BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Kongeriget Danmark

Patent application No.: PA 2001 00859

Date of filing: 31 May 2001

Applicant: (Name and address)
Novo Nordisk A/S
Novo Allé
DK-2880 Bagsværd
Denmark

Title: Hidtil ukendte forbindelser.

IPC: -

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

Patent- og Varemærkestyrelsen
Kongeriget Danmark

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

25 February 2004

Pia Petersen

PATENT- OG VAREMÆRKESTYRELSEN

HUMAN KUNITZ-TYPE PROTEASE INHIBITOR

FIELD OF THE INVENTION

The present invention relates to novel human kunitz-type protease inhibitors as well as polynucleotide constructs encoding such inhibitors, vectors and host cells comprising and expressing the inhibitors, pharmaceutical compositions, uses and methods of treatment.

BACKGROUND OF THE INVENTION

Polymorphonuclear leukocytes (neutrophils or PMNs) and mononuclear phagocytes (monocytes) play an important part in tissue injury, infection, acute and chronic inflammation and wound healing. The cells migrate from the blood to the site of inflammation and, following appropriate stimulation, they release oxidant compounds as well as granules containing a variety of proteolytic enzymes. The secretory granules contain, i.a., alkaline phosphatase, metalloproteinases such as gelatinase and collagenase and serine proteases such as neutrophil elastase, cathepsin G and proteinase 3.

In the azurophil granules of the leukocytes, the serine proteases neutrophil elastase, cathepsin G and proteinase-3 are packed as active enzymes complexed with glucosaminoglycans. These complexes are inactive but dissociate on secretion to release the active enzymes. To neutralize the protease activity, large amounts of the inhibitors α_1 -proteinase inhibitor (α_1-PI) and α_1 -chymotrypsin inhibitor (α_1 -ChI) are found in plasma. However, the PMNs are able to inactivate the inhibitors locally. Thus, α_1 -PI which is the most important inhibitor of neutrophil elastase is sensitive to oxidation at the reactive center (Met-358) by oxygen metabolites produced by triggered PMNs. This reduces the affinity of α_1 -PI for neutrophil elastase by approximately 2000 times.

After local neutralization of α_1 -PI, the elastase is able to degrade a number of inhibitors of other proteolytic enzymes. Elastase cleaves α_1 -ChI and thereby promotes cathepsin G activity. It also cleaves tissue inhibitor of metalloproteinase (TIMP), resulting in tissue degradation by metalloproteinases. Furthermore, elastase cleaves antithrombin III and heparin cofactor II, and tissue factor pathway inhibitor-1 (TFPI-1) which probably promotes clot formation. On the other hand, the ability of neutrophil elastase to degrade coagulation factors is assumed to have the opposite effect so that the total effect of elastase is unclear. The effect of neutrophil elastase on fibrinolysis is less ambiguous. Fibrinolytic activity increases when the elastase cleaves the plasminogen activator inhibitor and the alpha$_2$ plasmin inhibitor. Besides, both of these inhibitors are oxidated and inactivated by O_2 metabolites.

PMNs contain large quantities of serine proteases, and about 200 mg of each of the leukocyte proteases are released daily to deal with invasive agents in the body. Acute inflammation leads to a many-fold increase in the amount of enzyme released. Under normal conditions, proteolysis is kept at an acceptably low level by large amounts of the inhibitors α_1 -PI, α_1 -ChI and α_2 -macroglobulin. There is some indication, however, that a number of chronic diseases is caused by pathological proteolysis due to overstimulation of the PMNs, for instance caused by autoimmune response, chronic infection, tobacco smoke or other irritants, etc.

Aprotinin (bovine pancreatic trypsin inhibitor) is a kunitz-type inhibitor, known to inhibit various serine proteases, including trypsin, chymotrypsin, plasmin and kallikrein, and is used therapeutically in the treatment of acute pancreatitis, various states of shock syndrome, hyperfibrinolytic hemorrhage and myocardial infarction (ycf., for instance, J. E. Trapnell et al, Brit. J. Surg. 61, 1974, p. 177; J. McMichan et al., Circulatory shock 9, 1982, p. 107; L. M. Auer et al., Acta Neurochir. 49, 1979, p. 207; G. Sher, Am. J. Obstet. Gynecol. 129, 1977, p. 164; and B. Schneider, Artzneim.-Forsch. 26, 1976, p. 1606). Administration of aprotinin in high doses significantly reduces blood loss in connection with cardiac surgery, including cardiopulmonary bypass operations (cf., for instance, B. P. Bidstrup et al., J. Thorac. Cardiovasc. Surg. 97, 1989, pp. 364-372; W. van Oeveren et al., Ann. Thorac. Surg. 44, 1987, pp. 640-645). It has previously been demonstrated (cf. H. R. Wenzel and H. Tschesche, Angew. Chem. Internat. Ed. 20, 1981, p. 295) that certain aprotinin analogues, e.g. aprotinin(1-58, Val15) exhibit a relatively high selectivity for granulocyte elastase and an inhibitory effect on collagenase. Aprotinin (1-58, Ala15) has a weak effect on elastase, while aprotinin (3-58, Arg15, Ala17, Ser42) exhibits an excellent plasma kallikrein inhibitory effect (cf. WO 89/10374).

However, when administered in vivo, aprotinin has been found to have a nephrotoxic effect in rats, rabbits and dogs after repeated injections of relatively high doses of aprotinin (Bayer, Trasylol, Inhibitor of proteinase; E. Glaser et al. in "Verhandlungen der Deutschen Gesellschaft für Innere Medizin, 78. Kongress", Bergmann, M unchen, 1972, pp. 1612-1614). The nephrotoxicity (i.a. appearing in the form of lesions) observed for aprotinin might be ascribed to the accumulation of aprotinin in the proximal tubulus cells of the kidneys as a result of the high positive net charge of aprotinin which causes it to be bound to the negatively charged surfaces of the tubuli. This nephrotoxicity makes aprotinin less suitable for clinical purposes, in particular those requiring administration of large doses of the inhibitor (such as cardiopulmonary bypass operations). Besides, aprotinin is a bovine protein which may therefore contain one or more epitopes which may give rise to an undesirable immune response on administration of aprotinin to humans.

The Kunitz inhibitors, are generally basic, low molecular weight proteins comprising one or more inhibitory domains ("Kunitz domains"). The Kunitz domain is a folding domain of approximately 50-60 residues, which forms a central anti-parallel beta sheet and a short C-terminal helix. This characteristic domain comprises six cysteine residues that form three disulfide bonds, resulting in a double-loop structure. Between the N-terminal region and the first beta strand resides the active inhibitory binding loop. This binding loop is disulfide bonded through the P2 Cys residue to the hairpin loop formed between the last two beta strands. Isolated Kunitz domains from a variety of proteinase inhibitors have been shown to have inhibitory activity (e.g., Petersen et al., Eur. J. Biochem. 125:310-316, 1996; Wagner et al., Biochem. Biophys. Res. Comm. 186:1138-1145, 1992; Dennis et al., J. Biol. Chem. 270:25411-25417, 1995).

Human proteinase inhibitors comprising one or more Kunitz domains include TFPI-1, TFPI-2, inter α trypsin inhibitor (Iαl), amyloid β-protein precursor (AβPP), amyloid protein precursor homologue (APPH), placental bikunin , α3-chain of collagen type VI (CA3VI), α1-chain of collagen type VII (CA1VII), and the multi-domain protein, WFIKKN.

TFPI-1, an extrinsic pathway inhibitor and a natural anticoagulant, contains three tandemly linked Kunitz inhibitor domains. The amino-terminal Kunitz domain inhibits factor VIIa, plasmin, and cathepsin G; the second domain inhibits factor Xa, trypsin, and chymotrypsin; and the third domain has no known activity. TFPI-2 has been shown to be an inhibitor of the amidolytic and proteolytic activities of human factor VIIa-tissue factor complex, factor XIa, plasma kallikrein, and plasmin (Sprecher et al., Proc. Natl. Acad. Sci. USA 91:3353-3357, 1994; Petersen et al., Biochem. 35:266-272, 1996). The ability of TFPI-2 to inhibit the factor VIIa-tissue factor complex and its relatively high levels of transcription in umbilical vein endothelial cells, placenta and liver suggests a specialized role for this protein in hemostasis. Placental bikunin is a serine proteinase inhibitor containing two Kunitz domains (Delaria et al., J. Biol. Chem. 272:12209-12214, 1997). Individual Kunitz domains of bikunin have been expressed and shown to be potent inhibitors of trypsin, chymotrypsin, plasmin, factor XIa, and tissue and plasma kallikrein (Delaria et al., ibid.). AβPP, APPH, CA3VI, CA1VII and WFIKKN contain 1 Kunitz domain as part of the protein structure.

Known Kunitz-type inhibitors lack specificity and may have low potency. Lack of specificity can result in undesirable side effects, such as nephrotoxicity that occurs after repeated injections of high doses of aprotinin. These limitations may be overcome by identification of novel Kunitz-type inhibitors, which may exhibit a more desired inhibitor profile than traditional Kunitz-type inhibitors. Hence, there is a need in the art for additional Kunitz-type proteinase inhibitors.

DESCRIPTION OF THE INVENTION

The present invention relates to novel protease inhibitors comprising a Kunitz domain. This Kunitz domain comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein this sequence is at least 80% identical to residues 5 through 55 of wild type human HKI-18 (SEQ ID NO:1), is referred to herein as "HKI-18 polypeptide". The wild type human HKI-18 Kunitz domain has the amino acid sequence:

Tyr Pro Val Arg Cys Leu Leu Pro Ser Ala His Gly Ser Cys Ala Asp Trp Ala Ala Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe Trp Tyr Gly Gly Cys His Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser (SEQ ID NO:1).

The HKI-18 Kunitz domain comprises six cysteine residues that form three disulfide bonds, resulting in a double-loop structure. Referring to SEQ ID NO:1, disulfide bonds in the HKI-18 kunitz domain are formed by paired cysteine residues Cys5 - Cys55; Cys14 – Cys38; and Cys30 - Cys51. Referring to SEQ ID NO:2, disulfide bonds in the HKI-18 kunitz domain are formed by paired cysteine residues Cys1 - Cys51; Cys10 – Cys34; and Cys26 – Cys47.

In a first aspect, the present invention relates to an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein this sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1. In one embodiment the polypeptide sequence is at least 90% identical to residues 5 through 55 of SEQ ID NO:1. In a further embodiment the polypeptide sequence is at least 95% identical to residues 5 through 55 of SEQ ID NO:1.

The term "isolated", when applied to a polypeptide, denotes that the polypeptide has been removed from its natural genetic milieu and is thus free of other extraneous or unwanted coding sequences, and is in a form suitable for use within genetically engineered polypeptide production systems. An "isolated" polypeptide is found in a condition other than its native environment, such as apart from blood and animal tissue. In a preferred form, the isolated polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin. It is preferred to provide the polypeptides in a highly purified form, i.e. greater than 95% pure, more preferably greater than 99% pure. When used in this context, the term "isolated" is also intended to include the presence of the same polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

The term "a polypeptide" as used herein, means a molecule comprising a polymer of amino acid residues joined by peptide bonds, whether produced naturally or synthetically. The "polypeptide" may comprise one or more polypeptide chains. A "polypeptide" may also comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a polypeptide by the cell in which the polypeptide is produced, and will vary with the type of cell. Polypeptides are defined herein in terms of

their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nevertheless.

The term "variant", as used herein, is intended to designate polypeptides, wherein one or more amino acid residues of the parent polypeptide have been deleted and/or wherein one or more amino acid residues have been inserted in the parent polypeptide and/or wherein one or more amino acid residues have been added to the parent polypeptide. Such addition can take place either at the N-terminal end or at the C-terminal end of the parent polypeptide or both.

In a second aspect, the invention relates to an isolated polypeptide obtained by cultivation of a host cell comprising a polynucleotide construct encoding a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein this sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1 in an appropriate growth medium under conditions allowing expression of the polypeptide and recovering the polypeptide from the culture medium.

The term "a polynucleotide" denotes a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized in vitro, or prepared from a combination of natural and synthetic molecules. The length of a polynucleotide molecule is given herein in terms of nucleotides (abbreviated "nt") or base pairs (abbreviated "bp"). The term "nucleotides" is used for both single- and double-stranded molecules where the context permits. When the term is applied to double-stranded molecules it is used to denote overall length and will be understood to be equivalent to the term "base pairs". It will be recognized by those skilled in the art that the two strands of a double-stranded polynucleotide may differ slightly in length and that the ends thereof may be staggered as a result of enzymatic cleavage; thus all nucleotides within a double-stranded polynucleotide molecule may not be paired. Such unpaired ends will in general not exceed 20 nt in length.

The term "HKI-18 gene" as used herein, means a gene encoding a HKI-18 polypeptide.

The term "a host cell", as used herein, represent any cell, including hybrid cells, in which heterologous DNA can be expressed. Typical host cells includes, but are not limited to bacterial cells, insect cells, yeast cells, mammalian cells, including human cells, such as BHK, CHO, HEK, and COS cells. Examples of suitable yeasts cells include cells of *Saccharomyces* spp. or *Schizosaccharomyces* spp., in particular strains of *Saccharomyces cerevisiae* or *Saccharomyces kluyveri*. A preferred strain of *Saccharomyces cerevisiae* is the strain MT663 (*MATa/MATα pep4-3/pep4-3 HIS4/his4 tpi::LEU2/tpi::LEU2* Cir$^+$). Strain MT663 is deposited in the Deutsche Sammlung von Mikroorganismen und Zellkulturen with deposit number DSM 6278 in connection with filing of WO 92/11378.

In a third aspect, the invention relates to a polynucleotide construct encoding a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein this sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1. In one embodiment the polynucleotide construct is a vector. In a further embodiment the polynucleotide construct is a DNA vector. In a further embodiment the polynucleotide construct is a plasmid DNA vector.

The term "vector", as used herein, means any nucleic acid entity capable of the amplification in a host cell. Thus, the vector may be an autonomously replicating vector, i.e. a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated. The choice of vector will often depend on the host cell into which it is to be introduced. Vectors include, but are not limited to plasmid vectors, phage vectors, viruses or cosmid vectors. Vectors usually contain a replication origin and at least one selectable gene, i.e., a gene which encodes a product which is readily detectable or the presence of which is essential for cell growth.

In a further aspect, the invention relates to a host cell comprising a polynucleotide construct encoding a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein this sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1.

In a further aspect, the invention relates to a method for producing an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1, the method comprising cultivating a host cell comprising a polynucleotide construct encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1 in an appropriate growth medium under conditions allowing expression of the polypeptide and recovering the polypeptide from the culture medium.

In a further aspect, the invention relates to a composition comprising an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1.

In a further aspect, the invention relates to a pharmaceutical composition comprising an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1; and optionally, a pharmaceutically acceptable carrier.

In a further aspect, the invention relates to the use of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1 for the preparation of a medicament for the treatment of systemic inflammatory response syndrome (SIRS), disseminated intravascular coagulation (DIC), acute respiratory distress syndrome (ARDS), sepsis, ischemia / reperfusion injury, acute pancreatitis, trauma, shock syndrome, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery.

The term "treatment", as used herein, means the administration of an effective amount of a therapeutically active polypeptide of the invention with the purpose of preventing any symptoms or disease state to develop or with the purpose of curing or easing such symptoms or disease states already developed. The term "treatment" is thus meant to include prophylactic treatment.

In a further aspect, the invention relates to a method for the treatment of systemic inflammatory response syndrome, acute pancreatitis, shock syndrome, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery, the method comprising administering a therapeutically effective amount of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein the sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1; to a subject in need thereof.

The term "subject" as used herein is intended to mean any animal, in particular mammals, such as humans, and may, where appropriate, be used interchangeably with the term "patient".

In one embodiment of the invention, the host cell is a eukaryotic cell. In a further embodiment of the invention, the host cell is of mammalian origin. In a further embodiment of the invention, the host cell is a yeast cell. In a further embodiment of the invention, the host cell is a strain of *Saccharomyces cerevisiae.*

In one embodiment of the invention the isolated polypeptide comprises a kunitz domain. In a prefered embodiment of the invention the isolated polypeptide has one kunitz domain.

In a further embodiment of the invention the isolated polypeptide has proteinase inhibiting activity.

In a further embodiment of the invention the isolated polypeptide inhibits at least one of the proteases selected from the group consisting of chymotrypsin, elastase, cathepsin G, proteinase 3, plasmin, plasma kallikrein, glandular kallikrein and trypsin.

In a further embodiment of the invention the isolated polypeptide is from 51 to 81 amino acid residues in length.

In a further embodiment of the invention the isolated polypeptide is from 51 to 67 residues in length.

In a further embodiment of the invention, the Xaa5 of SEQ ID NO:2 is Pro.

In a further embodiment of the invention, the Xaa7 of SEQ ID NO:2 is Thr.

In a further embodiment of the invention, the Xaa9 of SEQ ID NO:2 is Pro.

In a further embodiment of the invention, the Xaa11 of SEQ ID NO:2 is Arg.

In a further embodiment of the invention, the Xaa11 of SEQ ID NO:2 is Lys.

In a further embodiment of the invention, the Xaa12 of SEQ ID NO:2 is Ala.

In a further embodiment of the invention, the Xaa13 of SEQ ID NO:2 is Arg.

In a further embodiment of the invention, the Xaa14 of SEQ ID NO:2 is Ile.

In a further embodiment of the invention, the Xaa15 of SEQ ID NO:2 is Ile.

In a further embodiment of the invention, the Xaa30 of SEQ ID NO:2 is Val.

In a further embodiment of the invention, the Xaa35 of SEQ ID NO:2 is Arg.

In a further embodiment of the invention the isolated polypeptide comprises residues 5 through 55 of SEQ ID NO:1. In a specific embodiment the isolated polypeptide is residues 5 through 55 of SEQ ID NO:1.

In a further embodiment of the invention the isolated polypeptide comprises residues 1 through 58 of SEQ ID NO:1. In a specific embodiment the isolated polypeptide is residues 1 through 55 of SEQ ID NO:1.

In a further embodiment of the invention the isolated polypeptide comprises the sequence of SEQ ID NO:4. In a specific embodiment the isolated polypeptide is SEQ ID NO:4.

In a further embodiment of the invention the isolated polypeptide comprises the sequence of SEQ ID NO:5. In a specific embodiment the isolated polypeptide is SEQ ID NO:5.

In a further embodiment of the invention the isolated polypeptide comprises the sequence of SEQ ID NO:6. In a specific embodiment the isolated polypeptide is SEQ ID NO:6.

In a further embodiment of the invention the isolated polypeptide comprises the sequence of SEQ ID NO:7. In a specific embodiment the isolated polypeptide is SEQ ID NO:7.

The level of identity between amino acid sequences can be determined using the "FASTA" similarity search algorithm of Pearson and Lipman (Proc. Natl. Acad. Sci. USA 85:2444, 1988) and Pearson (Meth. Enzymol. 183:63, 1990).

In the present specification, amino acid residues are represented using abbreviations, as indicated in table 1, approved by IUPAC-IUB Commission on Biochemical Nomenclature (CBN). With respect to amino acids and the like having isomers, those which are rep-

resented by the following abbreviations are in natural L-form. Further, the left and right ends of an amino acid sequence of a peptide are, respectively, the N- and C-termini unless otherwise specified.

Table 1: Abbreviations for amino acid residues:

Amino acid	Tree-letter code	One-letter code
Glycine	Gly	G
Proline	Pro	P
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Cysteine	Cys	C
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Tryptophan	Trp	W
Histidine	His	H
Lysine	Lys	K
Arginine	Arg	R
Glutamine	Gln	Q
Asparagine	Asn	N
Glutamic Acid	Glu	E
Aspartic Acid	Asp	D
Serine	Ser	S
Threonine	Thr	T

For the preparation of recombinant HKI-18 polypeptides, a cloned human wild-type polynucleotide sequence encoding HKI-18 is used. This sequence may be modified to encode a desired HKI-18 polypeptide.

The amino acid sequence alterations may be accomplished by a variety of techniques. Modification of the DNA sequence may be by site-specific mutagenesis. Techniques for site-specific mutagenesis are well known in the art and are described by, for example, Zoller and Smith (*DNA* 3:479-488, 1984). Thus, using the nucleotide and amino acid sequences of human wild-type HKI-18, one may introduce the alterations of choice.

The polypeptides of the present invention can also comprise non-naturally occurring amino acid residues. Non-naturally occurring amino acids include, without limitation, beta-alanine, desaminohistidine, trans-3-methylproline, 2,4-methanoproline, cis-4-hydroxyproline, trans-4-hydroxyproline, N-methylglycine, allo-threonine, methylthreonine, hydroxyethylcysteine, hydroxyethylhomocysteine, nitroglutamine, homoglutamine, pipecolic acid, thiazolidine carboxylic acid, dehydroproline, 3- and 4-methylproline, 3,3-dimethylproline, tert-leucine, norvaline, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally

occurring amino acid residues into polypeptides. For example, an in vitro system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell-free system comprising an E. coli S30 extract and commercially available enzymes and other reagents. Polypeptides are purified by chromatography. See, for example, Robertson et al., J. Am. Chem. Soc. 113:2722, 1991; Ellman et al., Methods Enzymol. 202:301, 1991; Chung et al., Science 259:806-9, 1993; and Chung et al., Proc. Natl. Acad. Sci. USA 90:10145-9, 1993). In a second method, translation is carried out in Xenopus oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs (Turcatti et al., J. Biol. Chem. 271:19991-8, 1996). Within a third method, E. coli cells are cultured in the absence of a natural amino acid that is to be replaced (e.g., phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (e.g., 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the polypeptide in place of its natural counterpart. See, Koide et al., Biochem. 33:7470-6, 1994. Naturally occurring amino acid residues can be converted to non-naturally occurring species by in vitro chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand the range of substitutions (Wynn and Richards, Protein Sci. 2:395-403, 1993).

As previously noted, polynucleotides encoding HKI-18 polypeptides provided by the present invention include DNA and RNA. Methods for preparing DNA and RNA are well known in the art. In general, RNA is isolated from a tissue or cell that produces large amounts of RNA encoding HKI-18 polypeptides. Such tissues and cells are identified by Northern blotting (Thomas, Proc. Natl. Acad. Sci. USA 77:5201, 1980), and include spinal cord, trachea, heart, colon, small intestine, and stomach. Total RNA can be prepared using guanidine-HCl extraction followed by isolation by centrifugation in a CsCl gradient (Chirgwin et al., Biochemistry 18:52-94, 1979). Poly $(A)^+$ RNA is prepared from total RNA using the method of Aviv and Leder (Proc. Natl. Acad. Sci. USA 69:1408-12, 1972). Complementary DNA (cDNA) is prepared from poly $(A)^+$ RNA using known methods. In the alternative, genomic DNA can be isolated. Polynucleotides encoding HKI-18 polypeptides are then identified and isolated by, for example, hybridization or PCR.

A full-length clone encoding a HKI-18 polypeptide can be obtained by conventional cloning procedures. Complementary DNA (cDNA) clones are preferred, although for some applications (e.g., expression in transgenic animals) it may be preferable to use a genomic clone, or to modify a cDNA clone to include at least one genomic intron. Methods for preparing cDNA and genomic clones are well known and within the level of ordinary skill in the art,

and include the use of the sequence disclosed herein, or parts thereof, for probing or priming a library. Expression libraries can be probed with antibodies to HKI-18 polypeptides, receptor fragments, or other specific binding partners.

The polynucleotides of the present invention can also be synthesized using automated equipment ("gene machines"). Gene synthesis methods are well known in the art. See, for example, Glick and Pasternak, Molecular Biotechnology, Principles & Applications of Recombinant DNA, ASM Press, Washington, D.C., 1994; Itakura et al., Annu. Rev. Biochem. 53: 323-356, 1984; and Climie et al., Proc. Natl. Acad. Sci. USA 87:633-637, 1990.

The polynucleotide sequences encoding HKI-18 polypeptides disclosed herein can be used to isolate counterpart polynucleotides from other species (orthologs). These orthologous polynucleotides can be used, inter alia, to prepare the respective orthologous polypeptides. These other species include, but are not limited to mammalian, avian, amphibian, reptile, fish, insect and other vertebrate and invertebrate species. Of particular interest are polynucleotide sequences encoding HKI-18 polypeptides and HKI-18 polypeptides from other mammalian species, including murine, porcine, ovine, bovine, canine, feline, equine, and other primate polypeptides. Orthologs of human HKI-18 polypeptides can be cloned using information and compositions provided by the present invention in combination with conventional cloning techniques. For example, a cDNA can be cloned using mRNA obtained from a tissue or cell type that expresses HKI-18 polypeptides as disclosed herein. Suitable sources of mRNA can be identified by probing Northern blots with probes designed from the sequences disclosed herein. A library is then prepared from mRNA of a positive tissue or cell line. A cDNA encoding HKI-18 polypeptides can then be isolated by a variety of methods, such as by probing with a complete or partial human cDNA or with one or more sets of degenerate probes based on the disclosed sequences. A cDNA can also be cloned using the polymerase chain reaction, or PCR (Mullis, U.S. Pat. No. 4,683,202), using primers designed from the representative human HKI-18 polypeptides sequence disclosed herein. Within an additional method, the cDNA library can be used to transform or transfect host cells, and expression of the cDNA of interest can be detected with an antibody to HKI-18 polypeptide. Similar techniques can also be applied to the isolation of genomic clones.

Those skilled in the art will recognize that the sequence disclosed in SEQ ID NO:3 represents a single allele of the polynucleotide encoding human HKI-18 polypeptide and that natural variation, including allelic variation and alternative splicing, is expected to occur. Allelic variants of this sequence can be cloned by probing cDNA or genomic libraries from different individuals according to standard procedures. Allelic variants of the DNA sequence shown in SEQ ID NO:3, including those containing silent mutations and those in which mutations result in amino acid sequence changes, are within the scope of the present invention,

as are polypeptides which are allelic variants of SEQ ID NO:1. cDNAs generated from alternatively spliced mRNAs, which retain the proteinase inhibiting activity of HKI-18 polypeptide are included within the scope of the present invention, as are polypeptides encoded by such cDNAs and mRNAs. Allelic variants and splice variants of these sequences can be cloned by probing cDNA or genomic libraries from different individuals or tissues according to standard procedures known in the art.

The DNA sequences encoding the HKI-18 polypeptide are usually inserted into a recombinant vector which may be any vector, which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

The vector is preferably an expression vector in which the DNA sequence encoding the HKI-18 polypeptide is operably linked to additional segments required for transcription of the DNA. In general, the expression vector is derived from plasmid or viral DNA, or may contain elements of both. The term, "operably linked" indicates that the segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in a promoter and proceeds through the DNA sequence coding for the polypeptide.

The promoter may be any DNA sequence, which shows transcriptional activity in the host cell of choice and may be derived from genes encoding polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the DNA encoding the HKI-18 polypeptide in mammalian cells are the SV40 promoter (Subramani et al., *Mol. Cell Biol.* 1 (1981), 854 -864), the MT-1 (metallothionein gene) promoter (Palmiter et al., *Science* 222 (1983), 809 - 814), the CMV promoter (Boshart et al., *Cell* 41:521-530, 1985) or the adenovirus 2 major late promoter (Kaufman and Sharp, *Mol. Cell. Biol*, 2:1304-1319, 1982).

An example of a suitable promoter for use in insect cells is the polyhedrin promoter (US 4,745,051; Vasuvedan et al., *FEBS Lett.* 311, (1992) 7 - 11), the P10 promoter (J.M. Vlak et al., *J. Gen. Virology* 69, 1988, pp. 765-776), the *Autographa californica* polyhedrosis virus basic protein promoter (EP 397 485), the baculovirus immediate early gene 1 promoter (US 5,155,037; US 5,162,222), or the baculovirus 39K delayed-early gene promoter (US 5,155,037; US 5,162,222).

Examples of suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem.* 255 (1980), 12073 - 12080; Alber and

Kawasaki, *J. Mol. Appl. Gen.* 1 (1982), 419 - 434) or alcohol dehydrogenase genes (Young et al., in *Genetic Engineering of Microorganisms for Chemicals* (Hollaender et al, eds.), Plenum Press, New York, 1982), or the TPI1 (US 4,599,311) or ADH2-4c (Russell et al., *Nature* 304 (1983), 652 - 654) promoters.

Examples of suitable promoters for use in filamentous fungus host cells are, for instance, the ADH3 promoter (McKnight et al., *The EMBO J.* 4 (1985), 2093 - 2099) or the tpiA promoter. Examples of other useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral α-amylase, *A. niger* acid stable α-amylase, *A. niger* or *A. awamori* glucoamylase (gluA), *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase. Preferred are the TAKA-amylase and gluA promoters. Suitable promoters are mentioned in, e.g. EP 238 023 and EP 383 779.

The DNA sequences encoding the HKI-18 polypeptide may also, if necessary, be operably connected to a suitable terminator, such as the human growth hormone terminator (Palmiter et al., *Science* 222, 1983, pp. 809-814) or the TPI1 (Alber and Kawasaki, *J. Mol. Appl. Gen.* 1, 1982, pp. 419-434) or ADH3 (McKnight et al., *The EMBO J.* 4, 1985, pp. 2093-2099) terminators. The vector may also contain a set of RNA splice sites located downstream from the promoter and upstream from the insertion site for the polynucleotide sequence encoding the HKI-18 polypeptide itself. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes. Also contained in the expression vectors is a polyadenylation signal located downstream of the insertion site. Particularly preferred polyadenylation signals include the early or late polyadenylation signal from SV40 (Kaufman and Sharp, ibid.), the polyadenylation signal from the adenovirus 5 Elb region, the human growth hormone gene terminator (DeNoto et al. *Nuc. Acids Res.* 9:3719-3730, 1981) or the polyadenylation signal from the HKI-18 gene. The expression vectors may also include a noncoding viral leader sequence, such as the adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites; and enhancer sequences, such as the SV40 enhancer.

The recombinant vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. An example of such a sequence (when the host cell is a mammalian cell) is the SV40 origin of replication.

When the host cell is a yeast cell, suitable sequences enabling the vector to replicate are the yeast plasmid 2μ replication genes REP 1-3 and origin of replication.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the gene coding for dihydrofolate reductase (DHFR) or the *Schizosaccharomyces pombe* TPI gene (described by P.R. Russell, *Gene* 40, 1985, pp. 125-130), or one which confers resistance to a drug, e.g. ampicillin, kanamycin,

tetracyclin, chloramphenicol, neomycin, hygromycin or methotrexate. For filamentous fungi, selectable markers include *amdS, pyrG, argB, niaD* or *sC.*

For secretion from yeast cells, the secretory signal sequence may encode any signal peptide, which ensures efficient direction of the expressed HKI-18 polypeptide into the secretory pathway of the cell. The signal peptide may be naturally occurring signal peptide, or a functional part thereof, or it may be a synthetic peptide. Suitable signal peptides have been found to be the α-factor signal peptide (cf. US 4,870,008), the signal peptide of mouse salivary amylase (cf. O. Hagenbuchle et al., *Nature* 289, 1981, pp. 643-646), a modified carboxypeptidase signal peptide (cf. L.A. Valls et al., *Cell* 48, 1987, pp. 887-897), the yeast BAR1 signal peptide (cf. WO 87/02670), or the yeast aspartic protease 3 (YAP3) signal peptide (cf. M. Egel-Mitani et al., *Yeast* 6, 1990, pp. 127-137).

For efficient secretion in yeast, a sequence encoding a leader peptide may also be inserted downstream of the signal sequence and upstream of the DNA sequence encoding the HKI-18 polypeptide. The function of the leader peptide is to allow the expressed peptide to be directed from the endoplasmic reticulum to the Golgi apparatus and further to a secretory vesicle for secretion into the culture medium (i.e. exportation of the HKI-18 polypeptide across the cell wall or at least through the cellular membrane into the periplasmic space of the yeast cell). The leader peptide may be the yeast α-factor leader (the use of which is described in e.g. US 4,546,082, US 4,870,008, EP 16 201, EP 123 294, EP 123 544 and EP 163 529). Alternatively, the leader peptide may be a synthetic leader peptide, which is to say a leader peptide not found in nature. Synthetic leader peptides may, for instance, be constructed as described in WO 89/02463 or WO 92/11378.

For use in filamentous fungi, the signal peptide may conveniently be derived from a gene encoding an *Aspergillus* sp. amylase or glucoamylase, a gene encoding a *Rhizomucor miehei* lipase or protease or a *Humicola lanuginosa* lipase. The signal peptide is preferably derived from a gene encoding *A. oryzae* TAKA amylase, *A. niger* neutral α-amylase, *A. niger* acid-stable amylase, or *A. niger* glucoamylase. Suitable signal peptides are disclosed in, e.g. EP 238 023 and EP 215 594.

For use in insect cells, the signal peptide may conveniently be derived from an insect gene (cf. WO 90/05783), such as the lepidopteran *Manduca sexta* adipokinetic hormone precursor signal peptide (cf. US 5,023,328).

The procedures used to ligate the DNA sequences coding for the HKI-18 polypeptide, the promoter and optionally the terminator and/or secretory signal sequence, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989).

Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid. If on the same plasmid, the selectable marker and the gene of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Pat. No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA," to the mixture that is introduced into the cells.

After the cells have taken up the DNA, they are grown in an appropriate growth medium, typically 1-2 days, to begin expressing the gene of interest. As used herein the term "appropriate growth medium" means a medium containing nutrients and other components required for the growth of cells and the expression of the HKI-18 polypeptide of interest. Media generally include a carbon source, a nitrogen source, essential amino acids, essential sugars, vitamins, salts, phospholipids, protein and growth factors. Drug selection is then applied to select for the growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an amplifiable selectable marker the drug concentration may be increased to select for an increased copy number of the cloned sequences, thereby increasing expression levels. Clones of stably transfected cells are then screened for expression of the HKI-18 polypeptide of interest.

HKI-18 polypeptides are tested for activity in protease inhibition assays, a variety of which are known in the art. Preferred assays include those measuring inhibition of trypsin, chymotrypsin, plasmin, cathepsin G, and human leukocyte elastase. See, for example, Petersen et al., Eur. J. Biochem. 235:310-316, 1996. In a typical procedure, the inhibitory activity of a test compound is measured by incubating the test compound with the proteinase, then adding an appropriate substrate, typically a chromogenic peptide substrate. See, for example, Norris et al. (Biol. Chem. Hoppe-Seyler 371:37-42, 1990). Briefly, various concentrations of the inhibitor are incubated in the presence of trypsin, plasmin, and plasma kallikrein in a low-salt buffer at pH 7.4, 25.degree. C. After 30 minutes, the residual enzymatic activity is measured by the addition of a chromogenic substrate (e.g., S2251 (D-Val-Leu-Lys-Nan) or S2302 (D-Pro-Phe-Arg-Nan), available from Kabi, Stockholm, Sweden) and a 30-minute incubation. Inhibition of enzyme activity is indicated by a decrease in absorbance at 405 nm or fluorescence Em at 460 nm. From the results, the apparent inhibition constant K_i is calculated. The inhibition of coagulation factors (e.g., factor VIIa, factor Xa) can be measured using chromogenic substrates or in conventional coagulation assays (e.g., clotting time of normal human plasma; Dennis et al., ibid.).

HKI-18 polypeptides can be tested in animal models of disease, particularly tumor models, models of fibrinolysis, and models of imbalance of hemostasis. Suitable models are

known in the art. For example, inhibition of tumor metastasis can be assessed in mice into which cancerous cells or tumor tissue have been introduced by implantation or injection (e.g., Brown, Advan. Enzyme Regul. 35:293-301, 1995; Conway et al., Clin. Exp. Metastasis 14:115-124, 1996). Effects on fibrinolysis can be measured in a rat model wherein the enzyme batroxobin and radiolabeled fibrinogen are administered to test animals. Inhibition of fibrinogen activation by a test compound is seen as a reduction in the circulating level of the label as compared to animals not receiving the test compound. See, Lenfors and Gustafsson, Semin. Thromb. Hemost. 22:335-342, 1996. The effect on various states of systemic inflammatory response syndrome (ARDS, DIC) can be tested in animals treated with iv injection of LPS, heat-killed and/or live bacteria to induce sepsis (eg Taylor et al. Blood 91, 1609-1615, 1998; Welty-Wolf et al. Am Respir Crit Care Med 158; 610-619; 1998).HKI-18 polypeptides can be delivered to test animals by injection or infusion, or can be produced in vivo by way of, for example, viral or naked DNA delivery systems or transgenic expression.

Exemplary viral delivery systems include adenovirus, herpesvirus, vaccinia virus and adeno-associated virus (AAV) . Adenovirus, a double-stranded DNA virus, is currently the best studied gene transfer vector for delivery of heterologous nucleic acid (for a review, see Becker et al., Meth. Cell Biol. 43:161-189, 1994; and Douglas and Curiel, Science & Medicine 4:44-53, 1997). The adenovirus system offers several advantages: adenovirus can (i) accommodate relatively large DNA inserts; (ii) be grown to high titer; (iii) infect a broad range of mammalian cell types; and (iv) be used with a large number of available vectors containing different promoters. Also, because adenoviruses are stable in the bloodstream, they can be administered by intravenous injection. By deleting portions of the adenovirus genome, larger inserts (up to 7 kb) of heterologous DNA can be accommodated. These inserts can be incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. In an exemplary system, the essential E1 gene is deleted from the viral vector, and the virus will not replicate unless the E1 gene is provided by the host cell (e.g., the human 293 cell line). When intravenously administered to intact animals, adenovirus primarily targets the liver. If the adenoviral delivery system has an E1 gene deletion, the virus cannot replicate in the host cells. However, the host's tissue (e.g., liver) will express and process (and, if a signal sequence is present, secrete) the heterologous polypeptide. Secreted polypeptides will enter the circulation in the highly vascularized liver, and effects on the infected animal can be determined.

An alternative method of gene delivery comprises removing cells from the body and introducing a vector into the cells as a naked DNA plasmid. The transformed cells are then re-implanted in the body. Naked DNA vectors are introduced into host cells by methods known in the art, including transfection, electroporation, microinjection, transduction, cell fu-

sion, DEAE dextran, calcium phosphate precipitation, use of a gene gun, or use of a DNA vector transporter. See, Wu et al., J. Biol. Chem. 263:14621-14624, 1988; Wu et al., J. Biol. Chem. 267:963-967, 1992; and Johnston and Tang, Meth. Cell Biol. 43:353-365, 1994.

Transgenic mice, engineered to express a HKI-18 gene, and mice that exhibit a complete absence of the function of the HKI-18 gene, referred to as "knockout mice" (Snouwaert et al., Science 257:1083, 1992), can also be generated (Lowell et al., Nature 366:740-742, 1993). These mice are employed to study the HKI-18 gene and the encoded polypeptide in an in vivo system. Transgenic mice are particularly useful for investigating the role of HKI-18 polypeptides in early development because they allow the identification of developmental abnormalities or blocks resulting from the over- or underexpression of a specific factor.

The HKI-18 polypeptides of the present invention, including full-length polypeptides, biologically active fragments, and fusion polypeptides can be produced in genetically engineered host cells according to conventional techniques. Suitable host cells are those cell types that can be transformed or transfected with exogenous DNA and grown in culture, and include bacteria, fungal cells, and cultured higher eukaryotic cells. Techniques for manipulating cloned DNA molecules and introducing exogenous DNA into a variety of host cells are disclosed by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, and Ausubel et al., eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., N.Y., 1987.

In general, a DNA sequence encoding a HKI-18 polypeptide is operably linked to other genetic elements required for its expression, generally including a transcription promoter and terminator, within an expression vector. The vector will also commonly contain one or more selectable markers and one or more origins of replication, although those skilled in the art will recognize that within certain systems selectable markers may be provided on separate vectors, and replication of the exogenous DNA may be provided by integration into the host cell genome. Selection of promoters, terminators, selectable markers, vectors and other elements is a matter of routine design within the level of ordinary skill in the art. Many such elements are described in the literature and are available through commercial suppliers.

To direct a HKI-18 polypeptide into the secretory pathway of a host cell, a secretory signal sequence (also known as a leader sequence, prepro sequence or pre sequence) is provided in the expression vector. The secretory signal sequence may be that of the HKI-18 polypeptide, or may be derived from another secreted polypeptide (e.g., t-PA; see, U.S. Pat. No. 5,641,655) or synthesized de novo. The secretory signal sequence is operably linked to the DNA sequence encoding a HKI-18 polypeptide, i.e., the two sequences are joined in the correct reading frame and positioned to direct the newly sythesized polypeptide into the sec-

retory pathway of the host cell. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the polypeptide of interest, although certain signal sequences may be positioned elsewhere in the DNA sequence of interest (see, e.g., Welch et al., U.S. Pat. No. 5,037,743; Holland et al., U.S. Pat. No. 5,143,830).

Cultured mammalian cells are suitable hosts for use within the present invention. Methods for introducing exogenous DNA into mammalian host cells include calcium phosphate-mediated transfection (Wigler et al., Cell 14:725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7:603, 1981: Graham and Van der Eb, Virology 52:456, 1973), electroporation (Neumann et al., EMBO J. 1:841-845, 1982), DEAE-dextran mediated transfection (Ausubel et al., ibid.), and liposome-mediated transfection (Hawley-Nelson et al., Focus 15:73, 1993; Ciccarone et al., Focus 15:80, 1993). The production of recombinant polypeptides in cultured mammalian cells is disclosed, for example, by Levinson et al., U.S. Pat. No. 4,713,339; Hagen et al., U.S. Pat. No. 4,784,950; Palmiter et al., U.S. Pat. No. 4,579,821; and Ringold, U.S. Pat. No. 4,656,134. Suitable cultured mammalian cells include the COS-1 (ATCC No. CRL 1650), COS-7 (ATCC No. CRL 1651), BHK (ATCC No. CRL 1632), BHK 570 (ATCC No. CRL 10314), 293 (ATCC No. CRL 1573; Graham et al., J. Gen. Virol. 36:59-72, 1977), Rat Hep I (Rat hepatoma; ATCC CRL 1600), Rat Hep II (Rat hepatoma; ATCC CRL 1548), TCMK (ATCC CCL 139), Human lung (ATCC HB 8065), NCTC 1469 (ATCC CCL 9.1), DUKX cells (Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA* 77:4216-4220, 1980) and Chinese hamster ovary (e.g. CHO-K1; ATCC No. CCL 61) cell lines.

Additional suitable cell lines are known in the art and available from public depositories such as the American Type Culture Collection, Rockville, Md. In general, strong transcription promoters are preferred, such as promoters from SV-40 or cytomegalovirus. See, e.g., U.S. Pat. No. 4,956,288. Other suitable promoters include those from metallothionein genes (U.S. Pat. Nos. 4,579,821 and 4,601,978) and the adenovirus major late promoter. Expression vectors for use in mammalian cells include pZP-1 and pZP-9, which have been deposited with the American Type Culture Collection, Rockville, Md. USA under accession numbers 98669 and 98668, respectively.

Drug selection is generally used to select for cultured mammalian cells into which foreign DNA has been inserted. Such cells are commonly referred to as "transfectants". Cells that have been cultured in the presence of the selective agent and are able to pass the gene of interest to their progeny are referred to as "stable transfectants." A preferred selectable marker is a gene encoding resistance to the antibiotic neomycin. Selection is carried out in the presence of a neomycin-type drug, such as G-418 or the like. Selection systems can also be used to increase the expression level of the gene of interest, a process referred to as "amplification." Amplification is carried out by culturing transfectants in the presence of a low

level of the selective agent and then increasing the amount of selective agent to select for cells that produce high levels of the products of the introduced genes. A preferred amplifiable selectable marker is dihydrofolate reductase, which confers resistance to methotrexate. Other drug resistance genes (e.g. hygromycin resistance, multi-drug resistance, puromycin acetyltransferase) can also be used.

Other higher eukaryotic cells can also be used as hosts, including insect cells, plant cells and avian cells. The use of Agrobacterium rhizogenes as a vector for expressing genes in plant cells has been reviewed by Sinkar et al., J. Biosci. (Bangalore) 11:47-58, 1987. Transformation of insect cells and production of foreign polypeptides therein is disclosed by Guarino et al., U.S. Pat. No. 5,162,222 and WIPO publication WO 94/06463.

Insect cells can be infected with recombinant baculovirus, commonly derived from Autographa californica nuclear polyhedrosis virus (AcNPV). See, King and Possee, The Baculovirus Expression System: A Laboratory Guide, London, Chapman & Hall; O'Reilly et al., Baculovirus Expression Vectors: A Laboratory Manual, New York, Oxford University Press., 1994; and Richardson, Ed., Baculovirus Expression Protocols. Methods in Molecular Biology, Humana Press, Totowa, N.J., 1995. Recombinant baculovirus can also be produced through the use of a transposon-based system described by Luckow et al. (J. Virol. 67:4566-4579, 1993). This system, which utilizes transfer vectors, is commercially available in kit form (Bac-to-Bac.TM. kit; Life Technologies, Rockville, Md.). The transfer vector (e.g., pFast-Bac1.TM.; Life Technologies) contains a Tn7 transposon to move the DNA encoding the polypeptide of interest into a baculovirus genome maintained in E. coli as a large plasmid called a "bacmid." See, Hill-Perkins and Possee, J. Gen. Virol. 71:971-976, 1990; Bonning et al., J. Gen. Virol. 75:1551-1556, 1994; and Chazenbalk and Rapoport, J. Biol. Chem. 270:1543-1549, 1995. In addition, transfer vectors can include an in-frame fusion with DNA encoding a polypeptide extension or affinity tag as disclosed above. Using techniques known in the art, a transfer vector containing a polynucleotide sequence encoding the HKI-18 polypeptide is transformed into E. coli host cells, and the cells are screened for bacmids which contain an interrupted lacZ gene indicative of recombinant baculovirus. The bacmid DNA containing the recombinant baculovirus genome is isolated, using common techniques, and used to transfect Spodoptera frugiperda cells, such as Sf9 cells. Recombinant virus that expresses HKI-18 polypeptide is subsequently produced. Recombinant viral stocks are made by methods commonly used the art.

For polypeptide production, the recombinant virus is used to infect host cells, typically a cell line derived from the fall armyworm, Spodoptera frugiperda (e.g., Sf9 or Sf21 cells) or Trichoplusia ni (e.g., High Five.TM. cells; Invitrogen, Carlsbad, Calif.). See, in general, Glick and Pasternak, Molecular Biotechnology: Principles and Applications of Recombi-

nant DNA, ASM Press, Washington, D.C., 1994. See also, U.S. Pat. No. 5,300,435. Serum-free media are used to grow and maintain the cells. Suitable media formulations are known in the art and can be obtained from commercial suppliers. The cells are grown up from an inoculation density of approximately 2-5 x10^5 cells to a density of 1-2 x 10^6 cells, at which time a recombinant viral stock is added at a multiplicity of infection (MOI) of 0.1 to 10, more typically near 3. Procedures used are generally described in available laboratory manuals (e.g., King and Possee, ibid.; O'Reilly et al., ibid.; Richardson, ibid.).

Fungal cells, including yeast cells, can also be used within the present invention. Yeast species of particular interest in this regard include *Saccharomyces cerevisiae, Saccharomyces kluyveri, Pichia pastoris*, and *Pichia methanolica*. Methods for transforming S. cerevisiae cells with exogenous DNA and producing recombinant polypeptides therefrom are disclosed by, for example, Kawasaki, U.S. Pat. No. 4,599,311; Kawasaki et al., U.S. Pat. No. 4,931,373; Brake, U.S. Pat. No. 4,870,008; Welch et al., U.S. Pat. No. 5,037,743; and Murray et al., U.S. Pat. No. 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (e.g., leucine). A preferred vector system for use in Saccharomyces cerevisiae is the POT1 vector system disclosed by Kawasaki et al. (U.S. Pat. No. 4,931,373), which allows transformed cells to be selected by growth in glucose-containing media. Suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, e.g., Kawasaki, U.S. Pat. No. 4,599,311; Kingsman et al., U.S. Pat. No. 4,615,974; and Bitter, U.S. Pat. No. 4,977,092) and alcohol dehydrogenase genes. See also U.S. Pats. Nos. 4,990,446; 5,063,154; 5,139,936 and 4,661,454.

Transformation systems for other yeasts, including Hansenula polymorpha, Schizosaccharomyces pombe, Kluyveromyces lactis, Kluyveromyces fragilis, Ustilago maydis, Pichia pastoris, Pichia methanolica, Pichia guillermondii and Candida maltosa are known in the art. See, for example, Gleeson et al., J. Gen. Microbiol. 132:3459-3465, 1986 and Cregg, U.S. Pat. No. 4,882,279. Aspergillus cells may be utilized according to the methods of McKnight et al., U.S. Pat. No. 4,935,349. Methods for transforming Acremonium chrysogenum are disclosed by Sumino et al., U.S. Pat. No. 5,162,228. Methods for transforming Neurospora are disclosed by Lambowitz, U.S. Pat. No. 4,486,533. The use of Pichia methanolica as host for the production of recombinant polypeptides is disclosed in U.S. Pats. No. 5,716,808, 5,736,383, 5,854,039, and 5,888,768; and WIPO Publications WO 97/17450 and W097/17451.

Prokaryotic host cells, including strains of the bacteria Escherichia coli, Bacillus and other genera are also useful host cells within the present invention. Techniques for transforming these hosts and expressing foreign DNA sequences cloned therein are well known in

the art (see, e.g., Sambrook et al., ibid.). When expressing a HKI-18 polypeptide in bacteria such as E. coli, the polypeptide may be retained in the cytoplasm, typically as insoluble granules, or may be directed to the periplasmic space by a bacterial secretion sequence. In the former case, the cells are lysed, and the granules are recovered and denatured using, for example, guanidine isothiocyanate or urea. The denatured polypeptide can then be refolded and dimerized by diluting the denaturant, such as by dialysis against a solution of urea and a combination of reduced and oxidized glutathione, followed by dialysis against a buffered saline solution. In the latter case, the polypeptide can be recovered from the periplasmic space in a soluble and functional form by disrupting the cells (by, for example, sonication or osmotic shock) to release the contents of the periplasmic space and recovering the polypeptide, thereby obviating the need for denaturation and refolding. Methods for producing heterologous disulfide bond-containing polypeptides in bacterial cells are disclosed by Georgiou et al., U.S. Pat. No. 6,083,715.

Transformed or transfected host cells are cultured according to conventional procedures in a culture medium containing nutrients and other components required for the growth of the chosen host cells. A variety of suitable media, including defined media and complex media, are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins and minerals. Media may also contain such components as growth factors or serum, as required. The growth medium will generally select for cells containing the exogenously added DNA by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker carried on the expression vector or co-transfected into the host cell.

HKI-18 polypeptides can also be prepared through chemical synthesis according to methods known in the art, including exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis. See, for example, Merrifield, J. Am. Chem. Soc. 85:2149, 1963; Stewart et al., Solid Phase Peptide Synthesis (2nd edition), Pierce Chemical Co., Rockford, Ill., 1984; Bayer and Rapp, Chem. Pept. Prot. 3:3, 1986; and Atherton et al., Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford, 1989.

It is preferred to purify the polypeptides of the present invention to about 80% purity, more preferably to about 90% purity, even more preferably about 95% purity, and particularly preferred is a pharmaceutically pure state, that is greater than 99.9% pure with respect to contaminating macromolecules, particularly other polypeptides and nucleic acids, and free of infectious and pyrogenic agents. Preferably, a purified polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin.

HKI-18 polypeptides are purified by conventional protein purification methods, typically by a combination of chromatographic techniques. Polypeptides comprising a polyhistidine affinity tag (typically about 6 histidine residues) are purified by affinity chromatography on a nickel chelate resin. See, for example, Houchuli et al., Bio/Technol. 6: 1321-1325, 1988.

Using methods known in the art, HKI-18 polypeptides can be produced glycosylated or non-glycosylated; pegylated or non-pegylated; and may or may not include an initial methionine amino acid residue.

The HKI-18 polypeptides are contemplated for use in the treatment or prevention of conditions associated with excessive proteinase activity, in particular an excess of trypsin, plasmin, kallikrein, elastase, cathepsin G, proteinase-3, thrombin, factor VIIa, factor IXa, factor Xa, factor XIa, factor XIIa, or matrix metalloproteinases. Such conditions include, but are not limited to, acute pancreatitis, cardiopulmonary bypass (CPB)-induced pulmonary injury, allergy-induced protease release, deep vein thrombosis, myocardial infarction, shock (including septic shock), hyperfibrinolytic hemorrhage, emphysema, rheumatoid arthritis, adult respiratory distress syndrome, chronic inflammatory bowel disease, psoriasis, systemic inflammatory response syndrome and other inflammatory conditions. HKI-18 polypeptides are also contemplated for use in preservation of platelet function, organ preservation, and wound healing.

HKI-18 polypeptides may be useful in the treatment of conditions arising from an imbalance in hemostasis, including acquired coagulopathies, primary fibrinolysis and fibrinolysis due to cirrhosis, and complications from high-dose thrombolytic therapy. Acquired coagulopathies can result from liver disease, uremia, acute disseminated intravascular coagulation, post-cardiopulmonary bypass, massive transfusion, or Warfarin overdose (Humphries, Transfusion Medicine 1:1181-1201, 1994). A deficiency or dysfunction in any of the procoagulant mechanisms predisposes the patient to either spontaneous hemorrhage or excess blood loss associated with trauma or surgery. Acquired coagulopathies usually involve a combination of deficiencies, such as deficiencies of a plurality of coagulation factors, and/or platelet dysfunction. In addition, patients with liver disease commonly experience increased fibrinolysis due to an inability to maintain normal levels of alpha 2 -antiplasmin and/or decreased hepatic clearance of plasminogen activators (Shuman, Hemorrhagic Disorders, in Bennet and Plum, eds. Cecil Textbook of Medicine, 20th ed., W. B. Saunders Co., 1996). Primary fibrinolysis results from a massive release of plasminogen activator. Conditions associated with primary fibrinolysis include carcinoma of the prostate, acute promyelocytic leukemia, hemangiomas, and sustained release of plasminogen activator by endothelial cells due to injection of venoms. The condition becomes critical when enough plasmin is activated to deplete the circulating level of $\alpha 2$ -antiplasmin (Shuman, ibid.). Data suggest that

plasmin on endothelial cells may be related to the pathophysiology of bleeding or rethrombosis observed in patients undergoing high-dose thrombolytic therapy for thrombosis. Plasmin may cause further damage to the thrombogenic surface of blood vessels after thrombolysis, which may result in rethrombosis (Okajima, J. Lab. Clin. Med. 126:1377-1384, 1995).

Additional antithrombotic uses of HKI-18 polypeptides include treatment or prevention of deep vein thrombosis, pulmonary embolism, and post-surgical thrombosis.

HKI-18 polypeptides may also be used within methods for inhibiting blood coagulation in mammals, such as in the treatment of disseminated intravascular coagulation. HKI-18 polypeptides may thus be used in place of known anticoagulants such as heparin, coumarin, and anti-thrombin III. Such methods will generally include administration of the polypeptide in an amount sufficient to produce a clinically significant inhibition of blood coagulation. Such amounts will vary with the nature of the condition to be treated, but can be predicted on the basis of known assays and experimental animal models, and will in general be within the ranges disclosed below.

HKI-18 polypeptides may also find therapeutic use in the blockage of proteolytic tissue degradation. Proteolysis of extracellular matrix, connective tissue, and other tissues and organs is an element of many diseases. This tissue destruction is beleived to be initiated when plasmin activates one or more matrix metalloproteinases (e.g., collagenase and metalloelastases). Inhibition of plasmin by HKI-18 polypeptides may thus be beneficial in the treatment of these conditions.

Matrix metalloproteinases (MMPs) are believed to play a role in metastases of cancers, abdominal aortic aneurysm, multiple sclerosis, rheumatoid arthritis, osteoarthritis, trauma and hemorrhagic shock, and cornial ulcers. MMPs produced by tumor cells break down and remodel tissue matrices during the process of metastatic spread. There is evidence to suggest that MMP inhibitors may block this activity (Brown, Advan. Enzyme Regul. 35:293-301, 1995). Abdominal aortic aneurysm is characterized by the degradation of extracellular matrix and loss of structural integrity of the aortic wall. Data suggest that plasmin may be important in the sequence of events leading to this destruction of aortic matrix (Jean-Claude et al., Surgery 116:472-478, 1994). Proteolytic enzymes are also believed to contribute to the inflammatory tissue damage of multiple sclerosis (Gijbels, J. Clin. Invest. 94:2177-2182, 1994). Rheumatoid arthritis is a chronic, systemic inflammatory disease predominantly affecting joints and other connective tissues, wherein proliferating inflammatory tissue (panus) may cause joint deformities and dysfunction (see, Arnett, in Cecil Textbook of Medicine, ibid.). Osteoarthritis is a chronic disease causing deterioration of the joint cartilage and other joint tissues and the formation of new bone (bone spurs) at the margins of the joints. There is evidence that MMPs participate in the degradation of collagen in the matrix of os-

teoarthritic articular cartilage. Inhition of MMPs results in the inhibition of the removal of collagen from cartilage matrix (Spirito, Inflam. Res. 44 (supp. 2):S131-S132, 1995; O'Byrne, Inflam. Res. 44 (supp. 2):S117-S118, 1995; Karran, Ann. Rheumatic Disease 54:662-669, 1995). HKI-18 polypeptides may also be useful in the treatment of trauma and hemorrhagic shock. Data suggest that administration of an MMP inhibitor after hemorrhage improves cardiovascular response, hepatocellular function, and microvascular blood flow in various organs (Wang, Shock 6:377-382, 1996). Corneal ulcers, which can result in blindness, manifest as a breakdown of the collagenous stromal tissue. Damage due to thermal or chemical injury to corneal surfaces often results in a chronic wound-healing situation. There is direct evidence for the role of MMPs in basement membrane defects associated with failure to re-epithelialize in cornea or skin (Fini, Am. J. Pathol. 149:1287-1302, 1996).

The HKI-18 polypeptides of the present invention may be combined with other therapeutic agents to augment the activity (e.g., antithrombotic or anticoagulant activity) of such agents. For example, a HKI-18 polypeptide may be used in combination with tissue plasminogen activator in thrombolytic therapy.

Doses of HKI-18 polypeptides will vary according to the severity of the condition being treated and may range from approximately 0.01 mg/kg to 10 mg/kg body weight, such as 0.1 mg/kg to 5 mg/kg, typically 0.1 mg/kg to 1 mg/kg. The polypeptides are typically formulated in a pharmaceutically acceptable carrier or vehicle. It is preferred to prepare them in a form suitable for injection or infusion, such as by dilution with sterile water, an isotonic saline or glucose solution, or similar vehicle. In the alternative, the polypeptide may be packaged as a lyophilized powder, optionally in combination with a pre-measured diluent, and resuspended immediately prior to use. Pharmaceutical compositions may further include one or more excipients, preservatives, solubilizers, buffering agents, albumin to prevent protein loss on vial surfaces, etc. Formulation methods are within the level of ordinary skill in the art. See, Remington: The Science and Practice of Pharmacy, Gennaro, ed., Mack Publishing Co., Easton, Pa., 19th ed., 1995.

Gene therapy provides an alternative therapeutic approach for delivery of HKI-18 polypeptides. If a mammal has a mutated or absent HKI-18 gene, a polynucleotide encoding a HKI-18 polypeptide can be introduced into the cells of the mammal. In one aspect the invention relates to method for introduction of a HKI-18 gene into a mammal. In one embodiment, a HKI-18 gene is introduced in vivo in a viral vector. Such vectors include an attenuated or defective DNA virus, such as herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. A defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to

cells in a specific, localized area, without concern that the vector can infect other cells. Examples of particular vectors include, without limitation, a defective herpes simplex virus 1 (HSV1) vector (Kaplitt et al., Molec. Cell. Neurosci. 2:320-30, 1991); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet et al., J. Clin. Invest. 90:626-30, 1992; and a defective adeno-associated virus vector (Samulski et al., J. Virol. 61:3096-101, 1987; Samulski et al., J. Virol. 63:3822-8, 1989). Within another embodiment, a HKI-18 gene can be introduced in a retroviral vector, as described, for example, by Anderson et al., U.S. Pat. No. 5,399,346; Mann et al. Cell 33:153, 1983; Temin et al., U.S. Pat. No. 4,650,764; Temin et al., U.S. Pat. No. 4,980,289; Markowitz et al., J. Virol. 62:1120, 1988; Temin et al., U.S. Pat. No. 5,124,263; Dougherty et al., WIPO Publication No. WO 95/07358; and Kuo et al., Blood 82:845, 1993. Alternatively, the vector can be introduced by lipofection in vivo using liposomes. Synthetic cationic lipids can be used to prepare liposomes for in vivo transfection of a gene encoding a marker (Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7, 1987; Mackey et al., Proc. Natl. Acad. Sci. USA 85:8027-31, 1988). Within a further embodiment, target cells are removed from the body, and a vector is introduced into the cells as a naked DNA plasmid. The transformed cells are then re-implanted into the body. Naked DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun or use of a DNA vector transporter. See, for example, Wu et al., J. Biol. Chem. 267:963-7, 1992; Wu et al., J. Biol. Chem. 263:14621-4, 1988.

HKI-18 polypeptides can also be used to prepare antibodies that specifically bind to HKI-18 polypeptides. As used herein, the term "antibodies" includes polyclonal antibodies, monoclonal antibodies, antigen-binding fragments thereof such as $F(ab')_2$ and Fab fragments, single chain antibodies, and the like, including genetically engineered antibodies. Non-human antibodies can be humanized by grafting only non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans is reduced. One skilled in the art can generate humanized antibodies with specific and different constant domains (i.e., different Ig subclasses) to facilitate or inhibit various immune functions associated with particular antibody constant domains. Alternative techniques for generating or selecting antibodies useful herein include in vitro exposure of lymphocytes to a HKI-

18 polypeptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled HKI-18 polypeptide). Antibodies are defined to be specifically binding if they bind to a HKI-18 polypeptide with an affinity at least 10-fold greater than the binding affinity to control (non-HKI-18) polypeptide. It is preferred that the antibodies exhibit a binding affinity (K_a) of 10^6 M^{-1} or greater, preferably 10^7 M^{-1} or greater, more preferably 10^8 M^{-1} or greater, and most preferably 10^9 M^{-1} or greater. The affinity of a monoclonal antibody can be readily determined by one of ordinary skill in the art (see, for example, Scatchard, Ann. NY Acad. Sci. 51: 660-672, 1949).

Methods for preparing polyclonal and monoclonal antibodies are well known in the art (see for example, Hurrell, J. G. R., Ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press, Inc., Boca Raton, Fla., 1982). As would be evident to one of ordinary skill in the art, polyclonal antibodies can be generated from a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, and rats. The immunogenicity of a HKI-18 polypeptide may be increased through the use of an adjuvant such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of a HKI-18 polypeptide or a portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof. If the polypeptide portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

Immunogenic fragments of HKI-18 polypeptides may be as small as 5 residues. It is preferred to use polypeptides that are hydrophilic or comprise a hydrophilic region. A variety of assays known to those skilled in the art can be utilized to detect antibodies that specifically bind to a HKI-18 polypeptide. Exemplary assays are described in detail in Antibodies: A Laboratory Manual, Harlow and Lane (Eds.), Cold Spring Harbor Laboratory Press, 1988. Representative examples of such assays include concurrent immunoelectrophoresis, radio-immunoassays, radio-immunoprecipitations, enzyme-linked immunosorbent assays (ELISA), dot blot assays, Western blot assays, inhibition or competition assays, and sandwich assays.

Antibodies to HKI-18 polypeptides may be used for affinity purification of HKI-18 polypeptides; within diagnostic assays for determining circulating levels of HKI-18 polypeptides; for detecting or quantitating soluble HKI-18 polypeptide as a marker of underlying pathology or disease; for immunolocalization within whole animals or tissue sections, including immunodiagnostic applications; for immunohistochemistry; for screening expression libraries; and for other uses that will be evident to those skilled in the art. For certain applications, including in vitro and in vivo diagnostic uses, it is advantageous to employ labelled antibodies.

Suitable direct tags or labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, chemiluminescent markers, magnetic particles and the like; indirect tags or labels may feature use of biotin-avidin or other complement/anti-complement pairs as intermediates.

HKI-18 polypeptides may be used in the laboratory or in commercial preparations of polypeptides from cultured cells. The polypeptides can be used alone to inhibit specific proteolysis or can be combined with other proteinase inhibitors to provide a "cocktail" with a broad spectrum of activity. Of particular interest is the inhibition of cellular proteases, which can be released during cell lysis. The polypeptides can also be used in the laboratory as a tissue culture additive to prevent cell detachment.

The present invention also provides reagents for use in diagnostic applications. For example, the HKI-18 gene, a probe comprising DNA or RNA or a subsequence thereof encoding the HKI-18 polypeptide can be used to determine if the HKI-18 gene is present on chromosome 7 or if a mutation has occurred. Detectable chromosomal aberrations at the locus of the HKI-18 gene include, but are not limited to, aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements. Such aberrations can be detected using polynucleotides of the present invention by employing molecular genetic techniques, such as restriction fragment length polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, and other genetic linkage analysis techniques known in the art (Sambrook et al., ibid.; Ausubel et. al., ibid.; Marian, Chest 108:255-65, 1995).

Polynucleotides encoding HKI-18 polypeptides can also be used for chromosomal mapping. Localization of the HKI-18 gene facilitates the establishment of directly proportional physical distances between newly discovered genes of interest and previously mapped markers, including the HKI-18 gene. The precise knowledge of a gene's position can be useful for a number of purposes, including: 1) determining if a newly identified sequence is part of an previously identified gene or gene segment and obtaining additional surrounding genetic sequences in various forms, such as YACs, BACs or cDNA clones; 2) providing a possible candidate gene for an inheritable disease which shows linkage to the same chromosomal region; and 3) cross-referencing model organisms, such as mouse, which may aid in determining what function a particular gene might have. A useful technique in this regard is radiation hybrid mapping, a somatic cell genetic technique developed for constructing high-resolution, contiguous maps of mammalian chromosomes (Cox et al., Science 250:245-50, 1990). Partial or full knowledge of a gene's sequence allows one to design PCR primers suitable for use with chromosomal radiation hybrid mapping panels. Radiation hybrid mapping panels, which are commercially available (e.g., the Stanford G3 RH Panel and the Gene-

Bridge 4 RH Panel, available from Research Genetics, Inc., Huntsville, Ala.), cover the entire human genome. These panels enable rapid, PCR-based chromosomal localizations and ordering of genes, sequence-tagged sites (STSs), and other nonpolymorphic and polymorphic markers within a region of interest.

Sequence tagged sites (STSs) can also be used independently for chromosomal localization. An STS is a DNA sequence that is unique in the human genome and can be used as a reference point for a particular chromosome or region of a chromosome. An STS is defined by a pair of oligonucleotide primers that are used in a polymerase chain reaction to specifically detect this site in the presence of all other genomic sequences. Since STSs are based solely on DNA sequence they can be completely described within an electronic database (for example, Database of Sequence Tagged Sites (dbSTS), GenBank; National Center for Biological Information, National Institutes of Health, Bethesda, Md.; http://www.ncbi.nlm.nih.gov) and can be searched with a gene sequence of interest for the mapping data contained within these short genomic landmark STS sequences.

BRIEF DESCRIPTION OF THE FIGURES

The present invention is described in further detail in the examples with reference to the appended drawings wherein

Fig. 1 shows the 174 bp DNA sequence encoding the human wild type HKI-18 kunitz domain.

Fig. 2 shows the 58 amino acid sequence of the human wild type HKI-18 kunitz domain.

Fig. 3 shows plasmid pMaUJ72 where the human wild type HKI-18 open reading frame has been cloned in pCANTAB 5E (Amersham Pharmacia) as described in example 1, "Cloning of human wild type HKI-18". Only restriction sites relevant for the construction of the plasmid described in the text have been indicated

Fig. 4 shows plasmid pMaUJ238. The plasmid contains an expression cassette comprising DNA encoding a fusion between the 212L leader and human wild type HKI-18. The plasmid was constructed as described in in example 1, "Construction of 212L-HKI18 fusion". Only restriction sites relevant for the construction of the plasmid described in the text have been indicated.

Fig. 5 shows an example on the construction of an expression cassette comprising DNA encoding a fusion between the 212L leader and human wild type HKI-18. PCR product 1 includes the 212L leader sequence and the Lys-Arg Kex2p cleavage site. PCR product 2 contains the human wild type HKI-18 open reading frame. PCR products 1 and 2 are used in a new PCR where the overlap extension ensures the resulting gene SOE product. The 5' end of primer b (corresponding to oMaUJ88) is complementary to the 3' end of primer c (corresponding to oMaUJ89). In the gene SOE reaction the two independent PCR products 1 and 2 are incubated with the primers a (corresponding to oMaUJ87) and d (corresponding to oMaUJ90).

Fig. 6 shows the pEA314 yeast plasmid. The plasmid contains an expression cassette comprising an *Eco*RI – *Xba*I fragment inserted between the transcription-promoter and the transcription-terminator of the *S. cerevisiae TPI* gene. *POT* is the selective marker, the Schizosaccharomyces pombe triosephosphate isomerase gene. Only restriction sites relevant for the description of pEA314 and for the construction of the 212L-HKI18 and alpha*L-HKI18 plasmids have been indicated.

Fig. 7 shows the nucleotide sequence and corresponding amino acid sequence of the 420 bp sequence *Eco*RI - *Xba*I encoding the 212L–HKI18 fusion polypeptide. The 212L leader sequence is underlined.

Fig. 8 shows 212L-HKI18-1 (212L-HKI18 P9, T11, K15, A16) and 212L-HKI18-2 (212L-HKI18 P9, T11, P13, K15, A16, R17, V34). The amino acid substitutions are shown with underlined and highlighted letters.

The present invention is further illustrated by the following examples which, however, are not to be construed as limiting the scope of protection. The features disclosed in the foregoing description and in the following examples may, both separately and in any combination thereof, be material for realizing the invention in diverse forms thereof.

EXAMPLES

Example 1

Cloning of human wild type HKI-18

Human wild type HKI-18 is included in a 133.5 kDa hypothetical polypeptide, which has been predicted from the DNA sequence EMBL:AF109907. The DNA sequence encoding the kunitz domain is seen in Figure 1, the amino acid sequence is seen in Figure 2.

The DNA encoding the human wild type HKI-18 domain was amplified by a polymerase chain reaction (PCR) in which human testis first-strand cDNA (Clontech, Multiple Tissue cDNA Panels) was used as template. 5 units Herculase Polymerase (Stratagene) and 1 nmol of the primers oMaUJ41 and oMaUJ42 (Table 3) were used in a 100 µl reaction. oMaUJ41 anneals to 19 bp in the 5´ end of the human wild type HKI-18 open reading frame and contains an *Sfi*I site in a 5´ extension. oMaUJ42 anneals to 22 bp in the 3´end of the human wild type HKI-18 open reading frame and contains an *Not*I site in a 5´extension. The PCR reaction was carried out as follows: 94 °C for 2 min, 4 cycles of 94 °C for 1 min; 60 °C 1 min; 72 °C for 30 sec, 4 cycles of 94 °C for 1 min; 55 °C 1 min; 72 °C for 30 sec; 25 cycles of 94 °C for 1 min; 50 °C 1 min; 72 °C for 30 sec. The resulting PCR product was digested with *Sfi*I/*Not*I and cloned in the *Sfi*I/*Not*I site of pCANTAB 5E (Amersham Pharmacia), resulting in pMaUJ72 (Figure 3). pMaUJ72 was checked for insert by appropriate restriction nucleases (e.g. *SFI*I and *Not*I) and was shown by sequence analysis using the primer oMaUJ11 (Table 3) to contain the proper sequence of human wild type HKI-18.

Construction of 212L-HKI18 fusion

The 212L-HKI18 construction, called pMaUJ238 (Figure 4), was created by gene splicing by overlap extension (gene SOE) (Horton *et al.*, Gene 77, 61-68, 1989), as illustrated in Figure 5. 5 units Taq polymerase (Promega) was used in 100 µl reactions. The primers and templates used in the construction of 212-HKI18 are listed in table 2 and described in the following.

pEA314 is identical to pKFN-1847 (Thim L. *et al.* FEBS 318, 345-352, 1993) except for a single nucleotide mutation changing Asn_{99} of human spasmolytic polypeptide (hSP) to Lys_{99}. pEA314 (Figure 6) contains an expression cassette comprising an *Eco*RI – *Xba*I fragment inserted between the transcription-promoter and the transcription-terminator of the *S. cerevisiae* TPI gene. The *Eco*RI – *Xba*I fragment encodes a fusion product composed of the

212L leader sequence, a Lys-Arg cleavage site for the dibasic processing endopeptidase KEX2, and the mutant hSP-Lys_{99}.

The 5' end of primer b (corresponding to oMaUJ88) is complementary to the 3' end of primer c (corresponding to oMaUJ89). In the gene SOE reaction (Figure 5) the two independent PCR products 1 and 2 are incubated with the primers a (corresponding to oMaUJ87) and d (corresponding to oMaUJ90). PCR product 1 includes the 212L leader sequence and the Lys-Arg Kex2p cleavage site. PCR product 2 contains the HKI18 ORF. The 5' end of oMaUJ87 contains a *Bst*XI site, the 5' end of oMaUJ90 contains an *Xba*I site. The resulting PCR product from the gene SOE reaction was digested with *Bst*XI and *Xba*I and ligated to pEA314 digested with *Bst*XI and *Xba*I. The final construct called pMaUJ238 (Figure 4) was propagated in *E. coli*, grown in the presence of ampicillin and isolated using standard techniques (Sambrook et al., *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory. Cold Spring Harbor, New York., 1989). pMaUJ238 was checked for insert by appropriate restriction nucleases (e.g. *Nco*I, *Bst*XI and *Xba*I) and was shown by sequence analysis using primers oMaUJ125 or oMaUJ126 (Table 3) to contain the proper construct of 212L-HKI18 (Figure 7).

Mutants of HKI-18 polypeptides

The mutants 212L-HKI18-1 (212L-HKI18 P9, T11, K15, A16) and 212L-HKI18-2 (212L-HKI18 P9, T11, P13, K15, A16, R17, V34) (Figure 8), were created by assembly PCR (Stemmer *et al.*, Gene 164, 49-53, 1995). The DNA's covering the mutated sequence were synthesized from oligo's 40-45 bp in length, using High Fidelity polymerase (Roche).

The PCR reaction was carried out as follows: 55 cycles of 94 °C in 30 sec; 52 °C in 30 sec; 72 °C 30 sec, followed by addition of the two outside primers oMaUJ224 and oMaUJ233 (Table 3) and 15 cycles of 94 °C in 30 sec; 50 °C in 30 sec; 72 °C 30 sec. The oligonucleotides used for assembly PCR introduce the four amino acid substitutions in HKI18-1 (oMaUJ224-233) and the 7 amino acid substitutions in HKI18-2 (oMaUJ224, 225, 228 and oMaUJ231-237), respectively. Furthermore, a silent mutation in oMaUJ224 removes the *Nco*I site, which is present in the 212L leader sequence.

The 5' end of oMaUJ224 contains an *Nco*I site, and the 5' end of oMaUJ233 contains an *Xba*I site. The resulting PCR products were digested with *Nco*I and *Xba*I and ligated into the *NcoI/Xba*I site of pEA314 (Figure 6) by partiel digestions. The resulting yeast plasmids pMaUJ365 and pMaUJ367 expressing HKI18-1 and HKI18-2, respectively, were propagated in *E. coli*, grown in the presence of ampicillin and isolated using standard techniques (Sambrook *et al.*, 1989). pMaUJ365 and pMaUJ367 were checked for insert by appropriate restriction nucleases (e.g. *Nco*I and *Xba*I) and are shown by sequence analysis using prim-

ers oMaUJ125 or oMaUJ126 (Table 3) to contain the proper constructs of 212L-HKI18-1 and 212L-HKI18-2, respectively.

Expression of HKI-18 polypeptides in yeast

The plasmids are transformed into *S. cerevisiae* strain MT663 (*MATa/MATα pep4-3/pep4-3 HIS4/his4 tpi::LEU2/tpi::LEU2* Cir$^+$). Strain MT663 was deposited in the Deutsche Sammlung von Mikroorganismen und Zellkulturen in connection with filing WO 92/11378 and was given the deposit number DSM 6278. Transformation of MT633 is conducted as described in WO 98/01535.

Yeast transformants harbouring pMaUJ238, pMaUJ365 and pMaUJ367, respectively, are selected by glucose utilization as carbon source on YPD (1% yeast extract, 2% peptone, 2% glucose) agar (2%) plates. The transformants yMaUJ33, yMaUJ69 and yMaUJ70 (Table 4) are selected for fermentation.

Yeast strains yMaUJ33, yMaUJ69 and yMaUJ70 are cultivated for 72 hours in ZYM media (2% yeast extract, 1% peptone, 6% glucose), and the cultures are pooled resulting in a final OD_{600} of approximately 15-20. After centrifugation the cell pellet is discarded and the supernatant is used for purification and characterization of the HKI-18 polypeptide.

Purification of HKI-18 polypeptides

The supernatant is adjusted to pH 7 and centrifuged to separate the cells. The supernatant is adjusted to pH 3 and centrifuged to a clarified fraction at 9000rpm, 50 minutes, 4°C on Sorwall SLA. This sample is used as application on an RPC-HPLC column (Jupiter C4 300Å 15µm). Solvents are 20mM citric acid, 50mM KCl pH 3,0, +/-60% EtOH. After application the sample is washed out with buffer A and an elution gradient (25%B - 80%B over 20cv) started. Fractions are collected and analysed by SDS for KPI-analogue presence.

Relevant fractions from the columns are collected and concentrated on a Centricon 3 (cut off 3kDa).

Finally the concentrated samples are buffer changed on PD10 column to a 20mM Na_2HPO_4, pH 7,5.

Characterization of HKI18

The concentrated sample is analysed by Mass-spectroscopy, SDS, Sequence analysis, Enzyme activity and HPLC.

Activity assay:

Inhibition of selected proteases by HKI-18 polypeptides are measured essentially as described by Petersen et al. FEBS Letters 338 53-57,1994. Chromogenic substrates H-D-Val-Leu-Lys-pNA; MeO-Suc-Arg-Pro-Tyr-pNA; H-D-Val-Leu-Arg-pNA; <Glu-Gly-Arg-pNA; H-D-Ile-Pro-Arg-pNA; <Glu-Pro-Arg-pNA; H-D-Pro-Phe-Arg-PNA; H-D-Phe-Pip-Arg-pNA; Cbo-D-Arg-Gly-Arg-pNA are purchased from Chromogenix (Moelndal, Sweden), MeO-CO-CHA-Gly-Arg-pNA from NycoMed (Oslo, Norway), and MeO-Suc-Ala-Ala-Pro-Val-pNA; Suc-Ala-Ala-Pro-Phe-pNA from Sigma (St.Louis, Mo, USA).

Trypsin, chymotrypsin, thrombin, plasmin, glandular kallikrein, human tissue-type plasminogen activator (tPA), activated protein C (APC) are from Sigma (St.Louis, Mo. USA) uPA are from FLUKA (Milwaukee, WI, USA). Recombinant factor VII (FVIIa) are from Novo Nordisk (Bagsværd, Denmark). Human leukocyte elastase (HLE) and cathepsin G (CatG) were purified by previously described procedures (Baugh, R.J. and Travis, J. (1976) Biochemistry 15 836-843). Activated factor X (FXa), activated factor XII (FXIIa) and recombinant tissue factor (TF) are from Calbiochem-Novabiochem Corporation (San Diego, Ca. USA).

Measurements of protease inhibition are performed by incubation with HKI-18 polypeptide for 30 min and subsequent substrate addition. Residual activity is then measured. The reaction takes place in microtiter wells in 100 mM NaCl, 50 mM TrisHCl, 0.01% Tween 80, pH 7.4 at 25°C in a total volume of 300 µl. Amidolytic activity is measured as the change in absorbance at 405 nm.

The apparent inhibition constant, K_i', are determined using the non-linear regression data analysis program Enzfitter^R (Biosoft, Cambridge, UK). K_i values are obtained by correcting for the effect of substrate according to the equation:

$$K_i = K_i'/(1+[S]/K_m)$$

SDS-analysis

Samples are diluted with sample buffer and applied on a 4-12% SDS gel in MES buffer. The gel is coomassie stained afterwards. One of the proteins in the MW standard is aprotinin.

HPLC-analysis

Purity is checked on an analytical column C4 Jupiter 5µm 4,6x250mm. Solvent system is TFA/CH3CN, flow 1,5ml/min. Gradient 5%B – 55%B over 20min. The purity measured by HPLC reflects the SDS pattern.

Table 2 shows the templates and primers used in the PCR's for construction of the plasmid pMaUJ72 and the yeast plasmid pMaUJ238.

Table 3 shows the DNA sequence of the primers used for construction and sequencing of the plasmids pMaUJ72, pMaUJ238, pMaUJ365 and pMaUJ367.

Table 4 shows yeast transformants harbouring pMaUJ238, pMaUJ365 and pMaUJ367, respectively.

Table 2:

Final construct	PCR reaction	Template	Upstream primer	Downstream primer
pMaUJ72		testis cDNA	oMaUJ41	oMaUJ42
pMaUJ238 (212L-HKI18)	1	pEA314	oMaUJ87	oMaUJ88
	2	pMaUJ72	oMaUJ89	oMaUJ90
	SOE	PCR product 1+2	oMaUJ87	oMaUJ90

Table 3:

Primer	Sequence
oMaUJ11	5' ACGCCAAGCTTTGGAGCC 3'
oMaUJ41	5' CCTTGATAGGCCCAGCCGGCCTACCCCGTGCGGTGCCTGC 3'
oMaUJ42	5' GGATGTCAAGCGGCCGCAGATCCCTGGCAG-CTGCTCATG 3'
oMaUJ87	5' CTGCAGAAGCACCATCAGGTTGGTG 3'
oMaUJ88	5' TCTCTTCTCCAATCTCTCAGCCATGGC 3'
oMaUJ89	5' CATGGCTGAGAGATTGGAGAAGAGATACCCCGTGCGG-TGCCTGCTGC 3'
oMaUJ90	5' CAGGCTGATCTAGACTTAAGATCCCTGGCAGCTGCTCA-TGCAC 3'
oMaUJ125	5' CAGGAATTCCATTCAAGAATAGTTC 3'
oMaUJ126	5' CCGTAGTCATCAATTTATTTTACATAACAC 3'
oMaUJ224	5' CTTTGGCTAACGTCGCCATGGCTGAGAGATTGGAGAAGAGATAC 3'
oMaUJ225	5' GGGCAGCAGGCACCGCACGGGGTATCTCTTCTCCAATCTCTC 3'
oMaUJ226	5' CCGTGCGGTGCCTGCTGCCCCCTGCCACTGGCTCTTGCAAAG 3'
oMaUJ227	5' GAAGTACCAGCGGGCAGCCCAGGCTTTGCAAGAGCCAGTGGCAG 3'
oMaUJ228	5' GGGCTGCCCGCTGGTACTTCGTTGCCTCTGTGGGCCAATGTAAC 3'
oMaUJ229	5' CATGACAGCCGCCATACCAGAAGCGGTTACATTGGCCCACAGAGG 3'
oMaUJ230	5' CTGGTATGGCGGCTGTCATGGCAATGCCAATAACTTTGCCTCGGAG 3'
oMaUJ231	5' CAGCTGCTCATGCACTCTTGCTCCGAGGCAAAGTTATTGG 3'
oMaUJ232	5' CAAGAGTGCATGAGCAGCTGCCAGGGATCTTAAGTCTAGA 3'
oMaUJ233	5' CACGGTCTTAGTTTCTAGACTTAAGATCCCTGG 3'
oMaUJ234	5' CCGTGCGGTGCCTGCTGCCCCCTGCCACTGGCCCTTGCAAAG 3'
oMaUJ235	5' GAAGTACCAGCGGGCAGCCCTGGCTTTGCAAGGGCCAGTGGCAG 3'
oMaUJ236	5' CATGACAGCCGCCATACACGAAGCGGTTACATTGGCCCACAGAGG 3'
oMaUJ237	5' CGTGTATGGCGGCTGTCATGGCAATGCCAATAACTTTGCCTCGGAG 3'

Table 4:

YEAST STRAIN	PLASMID	HKI18 ALLELE
yMaUJ33	pMaUJ238	HKI18
yMaUJ69	pMaUJ365	HKI18-1
yMaUJ70	pMaUJ367	HKI18-2

SEQUENCE LISTING:

SEQ ID NO.1 : Tyr Pro Val Arg Cys Leu Leu Pro Ser Ala His Gly Ser Cys Ala Asp Trp Ala Ala Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe Trp Tyr Gly Gly Cys His Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser

SEQ ID NO.2 : Cys Xaa2 Xaa3 Xaa4 Xaa5 Xaa6 Xaa7 Xaa8 Xaa9 Cys Xaa11 Xaa12 Xaa13 Xaa14 Xaa15 Xaa16 Xaa17 Xaa18 Xaa19 Xaa20 Xaa21 Xaa22 Xaa23 Xaa24 Xaa25 Cys Xaa27 Xaa28 Phe Xaa30 Xaa31 Xaa32 Gly Cys Xaa35 Xaa36 Xaa37 Xaa38 Asn Xaa40 Xaa41 Xaa42 Xaa43 Xaa44 Xaa45 Xaa46 Cys Xaa48 Xaa49 Xaa50 Cys

wherein Xaa2 is Ala, Val, Leu, Ser, Thr, Asn, Lys, Glu, Gln, Arg, Phe, Tyr, Met, or is absent;

wherein Xaa3 is Ala, Val, Leu, Ser, Thr, Asp, Glu, Gln, Phe, Met, or is absent;

wherein Xaa4 is Gly, Ala, Leu, Ser, Asp, Lys, Glu, Gln, Pro, or is absent;

wherein Xaa5 is Ala, Val, Leu, Glu, Ser, Asn, Lys, Glu, Tyr, Met, Pro, or is absent;

wherein Xaa6 is Ala, Val, Leu, Ser, Asp, Asn, Lys, Glu, Arg, Tyr, Met, or is absent;

wherein Xaa7 is Ala, Val, Thr, Asp, Lys, Glu, Gln, Arg, His, Tyr, Pro, or is absent;

wherein Xaa8 is Gly, Asp, or is absent;

wherein Xaa9 is Leu, Glu, Ser, Thr, Asn, Gln, Arg, Pro, or is absent;

wherein Xaa11 is Gly, Ala, Leu, Ser, Thr, Asn, Lys, Glu, Gln, Arg, Met, or is absent;

wherein Xaa12 is Gly, Ala, Thr, Asp, Glu, His, or is absent;

wherein Xaa13 is Leu, Glu, Ser, Asn, Glu, Arg, Phe, Trp, Tyr, Met, or is absent;

wherein Xaa14 is Ala, Val, Leu, Glu, Thr, Glu, Phe, Met, or is absent;

wherein Xaa15 is Ala, Val, Leu, Glu, Ser, Thr, Asn, Lys, Glu, Gln, Pro, or is absent;

wherein Xaa16 is Leu, Lys, Arg, His, or is absent;

wherein Xaa17 is Phe, Trp, Tyr, or is absent ;

wherein Xaa18 is Ala, His, Phe, Trp, Tyr, or is absent;

wherein Xaa19 is Phe, Tyr, or is absent;

wherein Xaa20 is Val, Ser, Asp, Asn, Arg, or is absent;

wherein Xaa21 is Gly, Ala, Leu, Glu, Ser, Asn, Lys, Phe, Pro, or is absent;

wherein Xaa22 is Val, Leu, Ser, Thr, Asn, Lys, Glu, Gln, Arg, Phe, Tyr, or is absent;

wherein Xaa23 is Ala, Val, Leu, Glu, Ser, Thr, Asp, Asn, Lys, Glu, Arg, Tyr, or is absent;

wherein Xaa24 is Gly, Asn, Lys, Glu, Gln, Arg, Tyr, Met, or is absent;
wherein Xaa25 is Ala, Leu, Glu, Ser, Thr, Lys, Glu, Gln, Arg, His, or is absent;
wherein Xaa27 is Ala, Val, Ser, Thr, Asp, Asn, Lys, Glu, Gln, Arg, His, or is absent;
wherein Xaa28 is Ala, Leu, Ser, Thr, Asn, Lys, Glu, Gln, Arg, Met, Pro, or is absent;
wherein Xaa30 is Ala, Val, Leu, Glu, Thr, Lys, Gln, Phe, Trp, Pro, or is absent;
wherein Xaa31 is Ser, Phe, Tyr, or is absent;
wherein Xaa32 is Gly, Ser, Thr, Arg, or is absent;
wherein Xaa35 is Gly, Leu, Asp, Asn, Glu, Gln, Arg, His, Tyr, Met, or is absent;
wherein Xaa36 is Gly, Ala, Arg, or is absent;
wherein Xaa37 is Ser, Asp, Asn, Lys, or is absent;
wherein Xaa38 is Gly, Ala, Ser, Asp, Asn, Lys, Glu, Gln, Arg, or is absent;
wherein Xaa40 is Ser, Asn, Lys, Arg, or is absent;
wherein Xaa41 is Phe, Tyr, or is absent;
wherein Xaa42 is Gly, Ala, Val, Leu, Thr, Asp, Asn, Lys, Glu, Gln, Arg, His, Tyr, Pro, or is absent;
wherein Xaa43 is Ser, Thr, Asp, Asn, Glu, Arg, or is absent;
wherein Xaa44 is Ala, Leu, Lys, Glu, Gln, Arg, Trp, or is absent
wherein Xaa45 is Ala, Asp, Lys, Glu, Gln, or is absent
wherein Xaa46 is Ala, Ser, Thr, Asp, Asn, Lys, Glu, Gln, Tyr, or is absent;
wherein Xaa48 is Leu, Ile, Glu, Asp, Lys, Glu, Gln, Arg, Met, or is absent;
wherein Xaa49 is Gly, Ala, Leu, Ser, Thr, Asp, Asn, Ly,s Glu, Gln, Arg, or is absent;
wherein Xaa50 is Ala, Ser, Thr, Va, Glu, Lys, Arg, Phe, Met, or is absent.

SEQ ID NO. 3:

```
1    TACCCCGTGC GGTGCCTGCT GCCCAGTGCC CATGGCTCTT GCGCAGACTG
     ATGGGGCACG CCACGGACGA CGGGTCACGG GTACCGAGAA CGCGTCTGAC
51   GGCTGCCCGC TGGTACTTCG TTGCCTCTGT GGGCCAATGT AACCGCTTCT
     CCGACGGGCG ACCATGAAGC AACGGAGACA CCCGGTTACA TTGGCGAAGA

101  GGTATGGCGG CTGCCATGGC AATGCCAATA ACTTTGCCTC GGAGCAAGAG
     CCATACCGCC GACGGTACCG TTACGGTTAT TGAAACGGAG CCTCGTTCTC
151  TGCATGAGCA GCTGCCAGGG ATCT
     ACGTACTCGT CGACGGTCCC TAGA
```

SEQ ID NO. 4: Tyr Pro Val Arg Cys Leu Leu Pro **Pro** Ala **Thr** Gly **Pro** Cys **Lys Ala Arg Ile Ile** Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe **Val** Tyr Gly Gly Cys **Arg** Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser

SEQ ID NO. 5: Tyr Pro Val Arg Cys Leu Leu Pro **Pro** Ala **Thr** Gly **Pro** Cys **Arg Ala Arg Ile Ile** Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe **Val** Tyr Gly Gly Cys **Arg** Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser

SEQ ID NO. 6: Tyr Pro Val Arg Cys Leu Leu Pro **Pro** Ala **Thr** Gly Ser Cys **Lys Ala** Trp Ala Ala Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe Trp Tyr Gly Gly Cys His Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser

SEQ ID NO. 7: Tyr Pro Val Arg Cys Leu Leu Pro **Pro** Ala **Thr** Gly **Pro** Cys **Lys Ala Arg** Ala Ala Arg Trp Tyr Phe Val Ala Ser Val Gly Gln Cys Asn Arg Phe **Val** Tyr Gly Gly Cys His Gly Asn Ala Asn Asn Phe Ala Ser Glu Gln Glu Cys Met Ser Ser Cys Gln Gly Ser

CLAIMS

1. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1.

2. The isolated polypeptide according to claim 1, wherein said polypeptide comprises a kunitz domain.

3. The isolated polypeptide according to any one of the claims 1-2, wherein said polypeptide has proteinase inhibiting activity.

4. The isolated polypeptide according to claim 3, which inhibits at least one of the proteases selected from the group consisting of chymotrypsin, elastase, cathepsin G, proteinase 3, plasmin, plasma kallikrein, glandular kallikrein and trypsin.

5. The isolated polypeptide according to any one of the claims 1-4, wherein said polypeptide is from 51 to 81 amino acid residues in length.

6. The isolated polypeptide according to any one of the claims 1-5, wherein said polypeptide is from 51 to 67 residues in length.

7. The isolated polypeptide according to any one of the claims 1-6, wherein Xaa5 of SEQ ID NO:2 is Pro.

8. The isolated polypeptide according to any one of the claims 1-7, wherein Xaa7 of SEQ ID NO:2 is Thr.

9. The isolated polypeptide according to any one of the claims 1-8, wherein Xaa9 of SEQ ID NO:2 is Pro.

10. The isolated polypeptide according to any one of the claims 1-9, wherein Xaa11 of SEQ ID NO:2 is Arg.

11. The isolated polypeptide according to any one of the claims 1-10, wherein Xaa11 of SEQ ID NO:2 is Lys.

12. The isolated polypeptide according to any one of the claims 1-11, wherein Xaa12 of SEQ ID NO:2 is Ala.

13. The isolated polypeptide according to any one of the claims 1-12, wherein Xaa13 of SEQ ID NO:2 Arg.

14. The isolated polypeptide according to any one of the claims 1-13, wherein Xaa14 of SEQ ID NO:2 is Ile.

15. The isolated polypeptide according to any one of the claims 1-14, wherein Xaa15 of SEQ ID NO:2 is Ile.

16. The isolated polypeptide according to any one of the claims 1-15, wherein Xaa30 of SEQ ID NO:2 is Val.

17. The isolated polypeptide according to any one of the claims 1-16, wherein Xaa35 of SEQ ID NO:2 is Arg.

18. The isolated polypeptide according to any one of the claims 1-6, wherein said sequence comprises residues 5 through 55 of SEQ ID NO:1.

19. The isolated polypeptide according to any one of the claims 1-6, wherein said sequence comprises residues 1 through 58 of SEQ ID NO:1.

20. The isolated polypeptide according to any one of the claims 1-6, wherein said sequence comprises a sequence independently selected from SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7.

21. An isolated polypeptide obtainable by cultivation of a host cell comprising a polynucleotide construct encoding a polypeptide according to any one of the claims 1-20 in an appropriate growth medium under conditions allowing expression of said polynucleotide construct and recovering said polypeptide from the culture medium.

22. The isolated polypeptide according to claim 21, wherein said host cell is a eukaryotic cell.

23. The isolated polypeptide according to claim 21-22, wherein said host cell is a mammalian cell.

24. The isolated polypeptide according to claim 21-22, wherein said host cell is a yeast cell.

25. The isolated polypeptide according to claim 24, wherein said host cell is a strain of *Saccharomyces cerevisiae.*

26. A polynucleotide construct encoding a polypeptide according to any one of the claims 1-20.

27. The polynucleotide construct according to claim 26, which is a vector.

28. A host cell comprising the polynucleotide construct according to any one of the claims 26-27.

29. The host cell according to claim 28, which is a eukaryotic cell.

30. The host cell according to claim 29, which is of mammalian origin.

31. The host cell according to claim 29, which is a yeast cell.

32. The host cell according to claim 31, which is a strain of *Saccharomyces cerevisiae.*

33. A method for producing the isolated polypeptide according to any one of claims 1-25, the method comprising cultivating a host cell as defined in any one of claims 28-32 in an appropriate growth medium under conditions allowing expression of said polynucleotide construct and recovering said polypeptide from the culture medium.

34. A composition comprising an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1.

35. A composition comprising an isolated polypeptide having the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to resi-

dues 5 through 55 of SEQ ID NO:1.

36. A composition comprising an isolated polypeptide according to any one of the claims 1-25.

37. A pharmaceutical composition comprising an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1; and optionally, a pharmaceutically acceptable carrier or vehicle.

38. A pharmaceutical composition comprising an isolated polypeptide according to any one of the claims 1-25; and optionally, a pharmaceutically acceptable carrier or vehicle.

39. Use of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1 for the preparation of a medicament for the treatment of systemic inflammatory response syndrome, acute pancreatitis, shock syndrome, disseminated intravascular coagulation, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery.

40. Use of an isolated polypeptide as defined in any one of the claims 1-25 for the preparation of a medicament for the treatment of systemic inflammatory response syndrome, acute pancreatitis, shock syndrome, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery.

41. A method for the treatment of systemic inflammatory response syndrome, acute pancreatitis, shock syndrome, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery, the method comprising administering a therapeutically or prophylactically effective amount of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or a variant thereof, wherein said sequence is at least 80% identical to residues 5 through 55 of SEQ ID NO:1; to a subject in need thereof.

42. A method for the treatment of systemic inflammatory response syndrome, acute pancreatitis, shock syndrome, hyperfibrinolytic hemorrhage, myocardial infarction or for prevention of blood loss during major surgery, the method comprising administering a therapeuti-

cally or prophylactically effective amount of an isolated polypeptide as defined in any one of claims 1-25 to a subject in need thereof.

5

ABSTRACT

The present invention relates to novel human kunitz-type protease inhibitors as well as polynucleotide constructs encoding such inhibitors, vectors and host cells comprising and expressing the inhibitors, pharmaceutical compositions, uses and methods of treatment.

FIGURES

FIGURE 1

```
  1  TACCCCGTGC GGTGCCTGCT GCCCAGTGCC CATGGCTCTT GCGCAGACTG
     ATGGGGCACG CCACGGACGA CGGGTCACGG GTACCGAGAA CGCGTCTGAC
 51  GGCTGCCCGC TGGTACTTCG TTGCCTCTGT GGGCCAATGT AACCGCTTCT
     CCGACGGGCG ACCATGAAGC AACGGAGACA CCCGGTTACA TTGGCGAAGA
101  GGTATGGCGG CTGCCATGGC AATGCCAATA ACTTTGCCTC GGAGCAAGAG
     CCATACCGCC GACGGTACCG TTACGGTTAT TGAAACGGAG CCTCGTTCTC
151  TGCATGAGCA GCTGCCAGGG ATCT
     ACGTACTCGT CGACGGTCCC TAGA
```

FIGURE 2

```
  1  YPVRCLLPSA HGSCADWAAR WYFVASVGQC NRFWYGGCHG NANNFASEQE
 51  CMSSCQGS
```

FIGURE 3

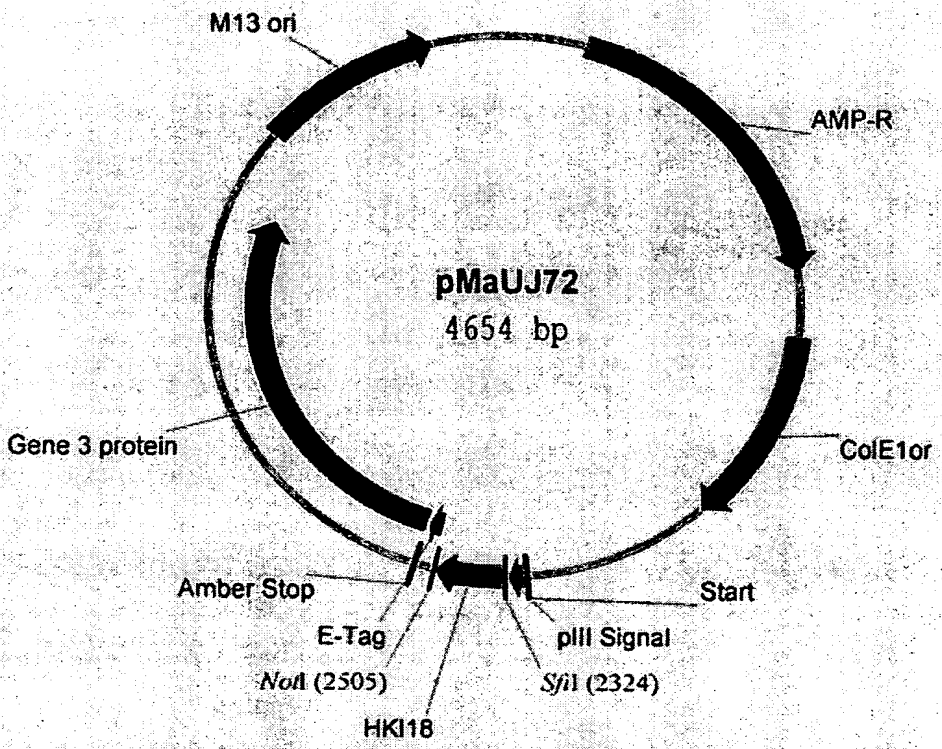

FIGURE 4

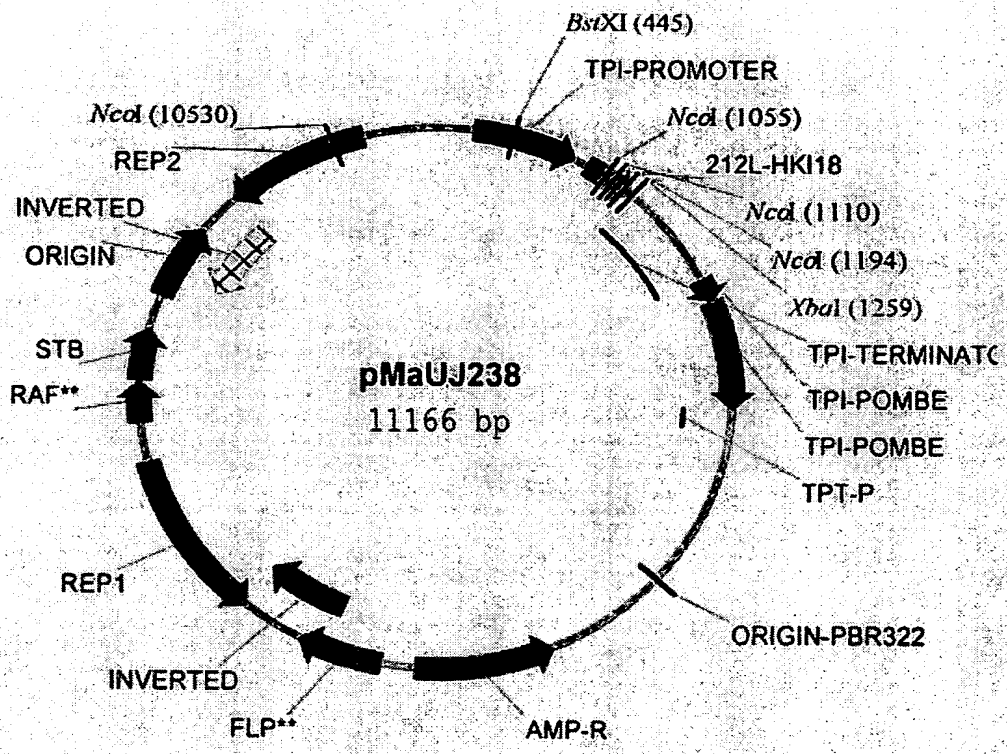

FIGURE 5

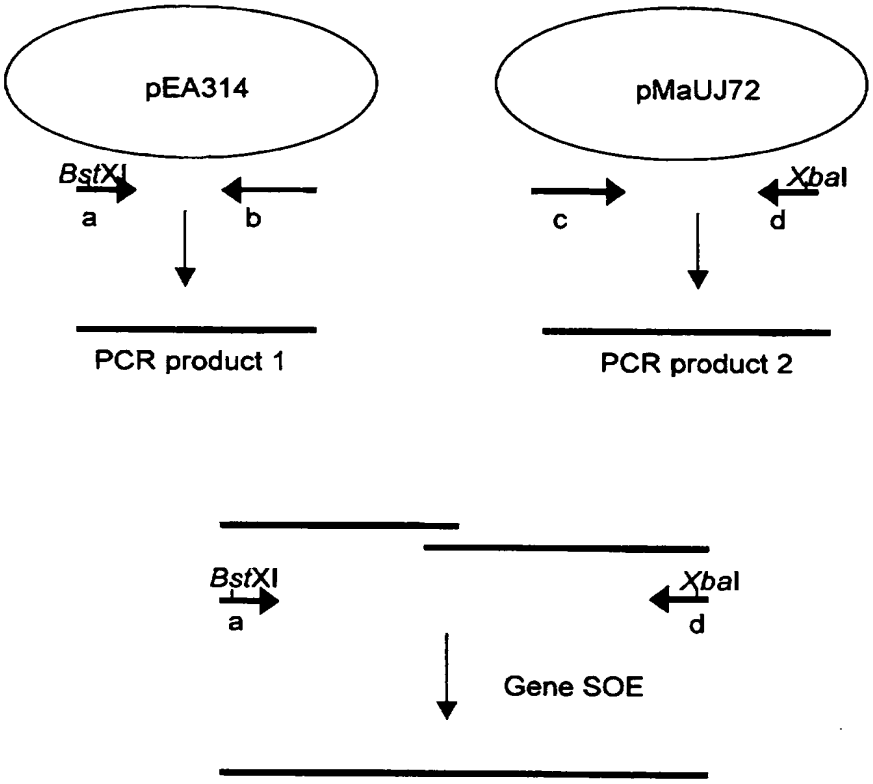
pEA314
pMaUJ72
BstXI
a
b
c
XbaI
d
PCR product 1
PCR product 2
BstXI
a
XbaI
d
Gene SOE

FIGURE 6

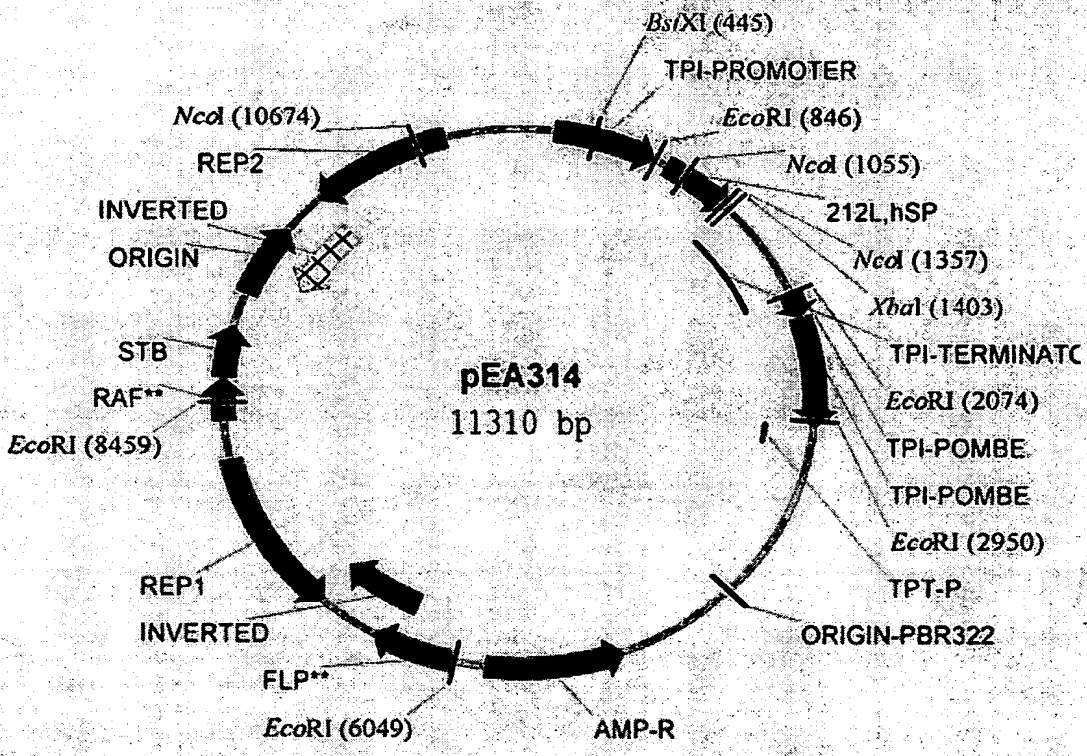

FIGURE 7

```
        EcoRI
        ~~~~~~
 845    GAATTCCATT CAAGAATAGT TCAAACAAGA AGATTACAAA CTATCAATTT CATACACAAT
                          M  K  A  V  F  L  V  L  S  L  I  G  F  C  W  ·
 905    ATAAACGACC AAAAGAATGA AGGCTGTTTT CTTGGTTTTG TCCTTGATCG GATTCTGCTG
        .  A  Q  P  V  T  G  D  E  S  S  V  E  I  P  E  E  S  L  I  I  ·
 965    GGCCCAACCA GTCACTGGCG ATGAATCATC TGTTGAGATT CCGGAAGAGT CTCTGATCAT
        .  A  E  N  T  T  L  A  N  V  A  M  A  E  R  L  E  K  R  Y  P  ·
1025    CGCTGAAAAC ACCACTTTGG CTAACGTCGC CATGGCTGAG AGATTGGAGA AGAGATACCC
        .  V  R  C  L  L  P  S  A  H  G  S  C  A  D  W  A  A  R  W  Y  ·
1085    CGTGCGGTGC CTGCTGCCCA GTGCCCATGG CTCTTGCGCA GACTGGGCTG CCCGCTGGTA
        .  F  V  A  S  V  G  Q  C  N  R  F  W  Y  G  G  C  H  G  N  A  ·
1145    CTTCGTTGCC TCTGTGGGCC AATGTAACCG CTTCTGGTAT GGCGGCTGCC ATGGCAATGC
                                                              XbaI
                                                              ~~~~~~
        .  N  N  F  A  S  E  Q  E  C  M  S  S  C  Q  G  S
1205    CAATAACTTT GCCTCGGAGC AAGAGTGCAT GAGCAGCTGC CAGGGATCTT AAGTCTAGA
```

FIGURE 8

212L-HKI18-1

```
   1  MKAVFLVLSL IGFCWAQPVT GDESSVEIPE ESLIIAENTT LANVAMAERL
  51  EKRYPVRCLL PPATGSCKAW AARWYFVASV GQCNRFWYGG CHGNANNFAS
 101  EQECMSSCQG S
```

212L-HKI18-2

```
   1  MKAVFLVLSL IGFCWAQPVT GDESSVEIPE ESLIIAENTT LANVAMAERL
  51  EKRYPVRCLL PPATGPCKAR AARWYFVASV GQCNRFVYGG CHGNANNFAS
 101  EQECMSSCQG S
```